PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68

A1

(11) International Publication Number: WO 99/02725

(43) International Publication Date: 21 January 1999 (21.01.99)

(21) International Application Number: PCT/GB98/02043

(22) International Filing Date: 13 July 1998 (13.07.98)

(30) Priority Data:

9714716.9	11 July 1997 (11.07.97)	GB
9719284.3	10 September 1997 (10.09.97)	GB
9726949.2	19 December 1997 (19.12.97)	GB

(71) Applicant *(for all designated States except US)*: BRAX GENOMICS LIMITED [GB/GB]; 13 Station Road, Cambridge CB1 2JB (GB).

(72) Inventors; and

(75) Inventors/Applicants *(for US only)*: SCHMIDT, Gunter [DE/GB]; Houghton Manor, Houghton, Cambs PE17 2BQ (GB). THOMPSON, Andrew, Hugin [GB/GB]; 25 Knoll Park, Alloway, Ayr KA7 4RH (GB).

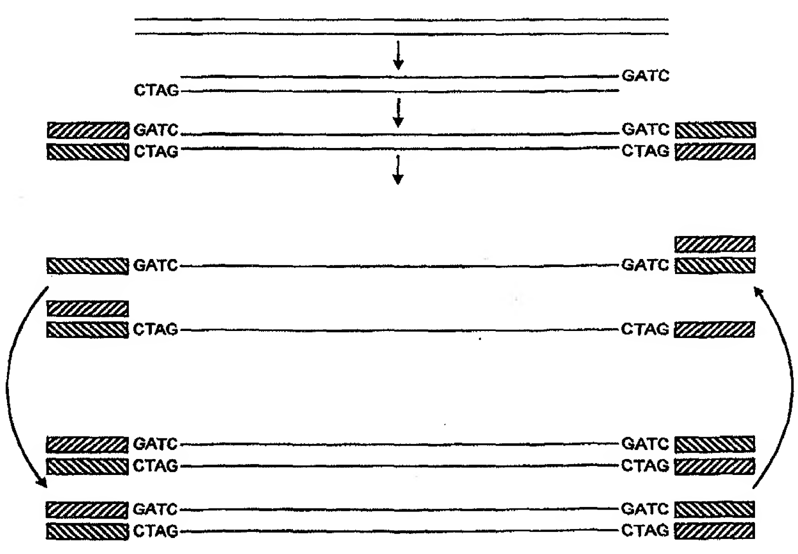
(74) Agents: DANIELS, Jeffrey, Nicholas et al.; Page White & Farrer, 54 Doughty Street, London WC1N 2LS (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published

With international search report.

(54) Title: CATEGORISING NUCLEIC ACID



(57) Abstract

Provided is a method for categorising nucleic acid, which method comprises producing a nucleic acid population by action of an endonuclease on double-stranded nucleic acid, such that each nucleic acid in the nucleic acid population has a double-stranded portion, contacting the nucleic acid population with one or more oligonucleotide sequences, and isolating nucleic acid which correctly hybridises to an oligonucleotide sequence, wherein each oligonucleotide sequence has a pre-determined recognition sequence, the nucleic acid being categorised by its ability to correctly hybridise to oligonucleotide sequences having the recognition sequence, the recognition sequence being situated such that it recognises a sequence in the double-stranded portion of the nucleic acid, one or more different recognition sequences being represented in the oligonucleotide sequences.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania
AM	Armenia
AT	Austria
AU	Australia
AZ	Azerbaijan
BA	Bosnia and Herzegovina
BB	Barbados
BE	Belgium
BF	Burkina Faso
BG	Bulgaria
BJ	Benin
BR	Brazil
BY	Belarus
CA	Canada
CF	Central African Republic
CG	Congo
CH	Switzerland
CI	Côte d'Ivoire
CM	Cameroon
CN	China
CU	Cuba
CZ	Czech Republic
DE	Germany
DK	Denmark
EE	Estonia
ES	Spain
FI	Finland
FR	France
GA	Gabon
GB	United Kingdom
GE	Georgia
GH	Ghana
GN	Guinea
GR	Greece
HU	Hungary
IE	Ireland
IL	Israel
IS	Iceland
IT	Italy
JP	Japan
KE	Kenya
KG	Kyrgyzstan
KP	Democratic People's Republic of Korea
KR	Republic of Korea
KZ	Kazakstan
LC	Saint Lucia
LI	Liechtenstein
LK	Sri Lanka
LR	Liberia
LS	Lesotho
LT	Lithuania
LU	Luxembourg
LV	Latvia
MC	Monaco
MD	Republic of Moldova
MG	Madagascar
MK	The former Yugoslav Republic of Macedonia
ML	Mali
MN	Mongolia
MR	Mauritania
MW	Malawi
MX	Mexico
NE	Niger
NL	Netherlands
NO	Norway
NZ	New Zealand
PL	Poland
PT	Portugal
RO	Romania
RU	Russian Federation
SD	Sudan
SE	Sweden
SG	Singapore
SI	Slovenia
SK	Slovakia
SN	Senegal
SZ	Swaziland
TD	Chad
TG	Togo
TJ	Tajikistan
TM	Turkmenistan
TR	Turkey
TT	Trinidad and Tobago
UA	Ukraine
UG	Uganda
US	United States of America
UZ	Uzbekistan
VN	Viet Nam
YU	Yugoslavia
ZW	Zimbabwe

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 98/02043

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WONG D M ET AL: "BRANCH CAPTURE REACTIONS: DISPLACERS DERIVED FROM ASYMMETRIC PCR" NUCLEIC ACIDS RESEARCH, vol. 19, no. 9, 11 May 1991, pages 2251-2259, XP000204316 see whole document and esp. figure 1 ---	1
X	GUILFOYLE R. A. ET AL.,: "Ligation-mediated PCR amplification of specific fragments from a class-II restriction endonuclease total digest" NUCLEIC ACIDS RESEARCH, vol. 25, no. 9, - 1 May 1997 pages 1854-1858, XP002076198	1-13, 25-36
Y	see the whole document ---	14-24
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 October 1998	29/10/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Müller, F

1

Form PCT/ISA/210 (second sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 98/02043

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 370 694 A (EASTMAN KODAK CO ;CETUS CORP (US)) 30 May 1990 see esp. claim 2 ---	14-24
A	EP 0 735 144 A (JAPAN RES DEV CORP) 2 October 1996 see the whole document ---	1-36
A	WO 94 01582 A (MEDICAL RES COUNCIL ;SIBSON DAVID ROSS (GB)) 20 January 1994 cited in the application see the whole document ---	1-36
A	US 5 508 169 A (DEUGAU KENNETH V ET AL) 16 April 1996 cited in the application see the whole document -----	1-36

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 98/02043

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0370694 A	30-05-1990	CA 2002076 A	21-05-1990
		DK 582189 A	22-05-1990
		JP 2299600 A	11-12-1990
EP 0735144 A	02-10-1996	JP 2763277 B	11-06-1998
		JP 9028399 A	04-02-1997
		JP 2763278 B	11-06-1998
		JP 8322598 A	10-12-1996
		AU 692685 B	11-06-1998
		AU 5031196 A	10-10-1996
		US 5707807 A	13-01-1998
WO 9401582 A	20-01-1994	AT 159986 T	15-11-1997
		AU 686563 B	12-02-1998
		AU 4575893 A	31-01-1994
		CA 2139944 A	20-01-1994
		DE 69315074 D	11-12-1997
		DE 69315074 T	05-03-1998
		EP 0650528 A	03-05-1995
		JP 7508883 T	05-10-1995
		US 5728524 A	17-03-1998
US 5508169 A	16-04-1996	CA 2036946 A	07-10-1991

PCT

WORLD INTELLECTUAL PROPERTY ORGANIZATION
International Bureau

INTERNATIONAL APPLICATION PUBLISHED UNDER THE PATENT COOPERATION TREATY (PCT)

(51) International Patent Classification 6 : C12Q 1/68	A1	(11) International Publication Number: WO 99/02725 (43) International Publication Date: 21 January 1999 (21.01.99)

(21) International Application Number: PCT/GB98/02043

(22) International Filing Date: 13 July 1998 (13.07.98)

(30) Priority Data:

9714716.9	11 July 1997 (11.07.97)	GB
9719284.3	10 September 1997 (10.09.97)	GB
9726949.2	19 December 1997 (19.12.97)	GB

(71) Applicant *(for all designated States except US)*: BRAX GENOMICS LIMITED [GB/GB]; 13 Station Road, Cambridge CB1 2JB (GB).

(72) Inventors; and

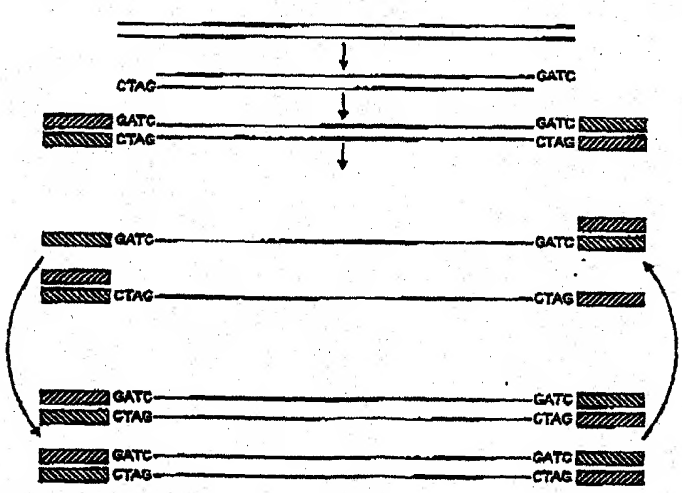
(75) Inventors/Applicants *(for US only)*: SCHMIDT, Gunter [DE/GB]; Houghton Manor, Houghton, Cambs PE17 2BQ (GB). THOMPSON, Andrew, Hugin [GB/GB]; 25 Knoll Park, Alloway, Ayr KA7 4RH (GB).

(74) Agents: DANIELS, Jeffrey, Nicholas et al.; Page White & Farrer, 54 Doughty Street, London WC1N 2LS (GB).

(81) Designated States: AL, AM, AT, AU, AZ, BA, BB, BG, BR, BY, CA, CH, CN, CU, CZ, DE, DK, EE, ES, FI, GB, GE, GH, GM, HR, HU, ID, IL, IS, JP, KE, KG, KP, KR, KZ, LC, LK, LR, LS, LT, LU, LV, MD, MG, MK, MN, MW, MX, NO, NZ, PL, PT, RO, RU, SD, SE, SG, SI, SK, SL, TJ, TM, TR, TT, UA, UG, US, UZ, VN, YU, ZW, ARIPO patent (GH, GM, KE, LS, MW, SD, SZ, UG, ZW), Eurasian patent (AM, AZ, BY, KG, KZ, MD, RU, TJ, TM), European patent (AT, BE, CH, CY, DE, DK, ES, FI, FR, GB, GR, IE, IT, LU, MC, NL, PT, SE), OAPI patent (BF, BJ, CF, CG, CI, CM, GA, GN, GW, ML, MR, NE, SN, TD, TG).

Published
With international search report.

(54) Title: CATEGORISING NUCLEIC ACID



(57) Abstract

Provided is a method for categorising nucleic acid, which method comprises producing a nucleic acid population by action of an endonuclease on double-stranded nucleic acid, such that each nucleic acid in the nucleic acid population has a double-stranded portion, contacting the nucleic acid population with one or more oligonucleotide sequences, and isolating nucleic acid which correctly hybridises to an oligonucleotide sequence, wherein each oligonucleotide sequence has a pre-determined recognition sequence, the nucleic acid being categorised by its ability to correctly hybridise to oligonucleotide sequences having the recognition sequence, the recognition sequence being situated such that it recognises a sequence in the double-stranded portion of the nucleic acid, one or more different recognition sequences being represented in the oligonucleotide sequences.

FOR THE PURPOSES OF INFORMATION ONLY

Codes used to identify States party to the PCT on the front pages of pamphlets publishing international applications under the PCT.

AL	Albania	ES	Spain	LS	Lesotho	SI	Slovenia
AM	Armenia	FI	Finland	LT	Lithuania	SK	Slovakia
AT	Austria	FR	France	LU	Luxembourg	SN	Senegal
AU	Australia	GA	Gabon	LV	Latvia	SZ	Swaziland
AZ	Azerbaijan	GB	United Kingdom	MC	Monaco	TD	Chad
BA	Bosnia and Herzegovina	GE	Georgia	MD	Republic of Moldova	TG	Togo
BB	Barbados	GH	Ghana	MG	Madagascar	TJ	Tajikistan
BE	Belgium	GN	Guinea	MK	The former Yugoslav Republic of Macedonia	TM	Turkmenistan
BF	Burkina Faso	GR	Greece			TR	Turkey
BG	Bulgaria	HU	Hungary	ML	Mali	TT	Trinidad and Tobago
BJ	Benin	IE	Ireland	MN	Mongolia	UA	Ukraine
BR	Brazil	IL	Israel	MR	Mauritania	UG	Uganda
BY	Belarus	IS	Iceland	MW	Malawi	US	United States of America
CA	Canada	IT	Italy	MX	Mexico	UZ	Uzbekistan
CF	Central African Republic	JP	Japan	NE	Niger	VN	Viet Nam
CG	Congo	KE	Kenya	NL	Netherlands	YU	Yugoslavia
CH	Switzerland	KG	Kyrgyzstan	NO	Norway	ZW	Zimbabwe
CI	Côte d'Ivoire	KP	Democratic People's Republic of Korea	NZ	New Zealand		
CM	Cameroon			PL	Poland		
CN	China	KR	Republic of Korea	PT	Portugal		
CU	Cuba	KZ	Kazakstan	RO	Romania		
CZ	Czech Republic	LC	Saint Lucia	RU	Russian Federation		
DE	Germany	LI	Liechtenstein	SD	Sudan		
DK	Denmark	LK	Sri Lanka	SE	Sweden		
EE	Estonia	LR	Liberia	SG	Singapore		

INTERNATIONAL SEARCH REPORT

Inter   .nal Application No
PCT/GB 98/02043

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WONG D M ET AL: "BRANCH CAPTURE REACTIONS: DISPLACERS DERIVED FROM ASYMMETRIC PCR" NUCLEIC ACIDS RESEARCH, vol. 19, no. 9, 11 May 1991, pages 2251-2259, XP000204316 see whole document and esp. figure 1	1
X	GUILFOYLE R. A. ET AL.,: "Ligation-mediated PCR amplification of specific fragments from a class-II restriction endonuclease total digest" NUCLEIC ACIDS RESEARCH, vol. 25, no. 9, - 1 May 1997 pages 1854-1858, XP002076198	1-13, 25-36
Y	see the whole document	14-24

-/--

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

* Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 October 1998	29/10/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Müller, F

1

Form PCT/ISA/210 (second sheet) (July 1992)

Intc. ...onal Application No
PCT/GB 98/02043

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category *	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 370 694 A (EASTMAN KODAK CO ;CETUS CORP (US)) 30 May 1990 see esp. claim 2	14-24
A	EP 0 735 144 A (JAPAN RES DEV CORP) 2 October 1996 see the whole document	1-36
A	WO 94 01582 A (MEDICAL RES COUNCIL ;SIBSON DAVID ROSS (GB)) 20 January 1994 cited in the application see the whole document	1-36
A	US 5 508 169 A (DEUGAU KENNETH V ET AL) 16 April 1996 cited in the application see the whole document	1-36

1

Information on patent family members

PCT/GB 98/02043

Patent document cited in search report		Publication date	Patent family member(s)			Publication date
EP 0370694	A	30-05-1990	CA	2002076	A	21-05-1990
			DK	582189	A	22-05-1990
			JP	2299600	A	11-12-1990
EP 0735144	A	02-10-1996	JP	2763277	B	11-06-1998
			JP	9028399	A	04-02-1997
			JP	2763278	B	11-06-1998
			JP	8322598	A	10-12-1996
			AU	692685	B	11-06-1998
			AU	5031196	A	10-10-1996
			US	5707807	A	13-01-1998
WO 9401582	A	20-01-1994	AT	159986	T	15-11-1997
			AU	686563	B	12-02-1998
			AU	4575893	A	31-01-1994
			CA	2139944	A	20-01-1994
			DE	69315074	D	11-12-1997
			DE	69315074	T	05-03-1998
			EP	0650528	A	03-05-1995
			JP	7508883	T	05-10-1995
			US	5728524	A	17-03-1998
US 5508169	A	16-04-1996	CA	2036946	A	07-10-1991

5/PRTS
514 Rec'd PCT/PTO 09/462635 11 JAN 2000

CATEGORISING NUCLEIC ACID

The present invention concerns a method for categorising nucleic acid. In particular, the invention concerns a method for sorting nucleic acid, which method permits reduction in the complexity of a nucleic acid population of approximately one order of magnitude, or more. The invention also relates to a kit for carrying out the above method.

Analysis of nucleic acids is fundamental to much of modern molecular biology. A particular feature of nucleic acids derived from living organism is that they are almost invariably complex populations of sequences present in widely varying quantities. In order to characterise these populations of nucleic acids it is usual to attempt to reduce the complexity of the population of nucleic acids in some way. Traditionally the approach has been to clone complex nucleic acid molecules into vectors to allow them to be isolated and either sub-cloned further or analysed directly. Cloning requires the use of biological hosts and these are often difficult to use and require a great deal of specialist knowledge for the cloning procedures to be successful. The traditional processes of cloning to generate libraries of sequences are also only partially automatable.

A problem which cloning does not address is how to isolate sequences which are present only at low copies in backgrounds of sequences present at high copy numbers. Various techniques have been developed to 'normalise' complex nucleic acid populations prior to cloning in order to increase the quantities of sequences at low copy numbers relative to those at high copy numbers. Subtractive hybridisation methods have been used to try and normalise cDNA populations.

PCT/GB93/01452 describes methods of molecular sorting which uses restriction endonucleases that generate ambiguous sticky-ends in the nucleic acid sample to be sorted. Adapters are designed with sticky ends complementary to a single sticky-end sequence or a subset of the these ambiguous sticky ends such that the individual sticky end or subset thereof is coupled to a distinct sequence in the double stranded region of the adapter. This allows subsets of the

adaptored nucleic acid to be amplified using specific primers corresponding to sequences within the adapter which in turn relate to the sequence of the sticky end of the adapter. US patent 5,508,169 (issued November 7, 1995) describes methods very similar to those disclosed in PCT/GB93/01452.

A problem with the above method is that the nucleic acids can be sorted only according to the sequence present on the sticky-ends of the nucleic acid. The sticky-end sequence is of limited length, as determined by the choice of restriction enzyme, thus the basis for sorting is limited.

It is an object of the present invention to provide a method which overcomes the above problems, and provides a wider basis on which sorting of nucleic acid populations can be carried out, not limited by the sticky-end sequence. It is also an object of this invention to provide methods to reduce the complexity of nucleic acid populations by allowing them to be sorted into sub-populations without cloning and to permit normalisation of these populations. This invention describes methods of sorting nucleic acid molecules that have a variety of applications including gene expression profiling, preparation of templates for sequencing, linkage analysis, etc. This invention provides methods of generating sorted libraries. In many applications it is preferable that these sorted nucleic acids be captured on a solid phase support.

Accordingly, the present invention provides a method for categorising nucleic acid, which method comprises producing a nucleic acid population by action of an endonuclease on double-stranded nucleic acid, such that each nucleic acid in the nucleic acid population has a double-stranded portion, contacting the nucleic acid population with one or more oligonucleotide sequences, and isolating nucleic acid which correctly hybridises to an oligonucleotide sequence, wherein each oligonucleotide sequence has a pre-determined recognition sequence, the nucleic acid being categorised by its ability to correctly hybridise to oligonucleotide sequences having the recognition sequence, the recognition sequence being situated such that it recognises a sequence in the double-stranded portion of the nucleic acid,

one or more different recognition sequences being represented in the oligonucleotide sequences.

The present invention also provides kit for categorising a nucleic acid, comprising one or more adaptors and one or more sets of oligonucleotide sequences, wherein the adaptors comprise nucleic acid having a double-stranded primer portion of a known sequence and a single-stranded portion of a pre-determined length, either each single-stranded portion of each nucleic acid in the adaptors having the same pre-determined sequence or all possible sequences of the single-stranded portion being represented in the adaptors, and wherein each oligonucleotide sequence comprises a first sequence, a second sequence attached to the first sequence and a third sequence attached to the second sequence, in which the first sequence is complementary to the sequence of the primer portion of the adaptor, the second sequence is the same sequence as the single-stranded portion of the adaptors or all possible second sequences of the same length as the single-stranded portion of the adaptors are represented within the set of oligonucleotides, and the third sequence comprises a pre-determined recognition sequence.

The invention will now be described in further detail by way of example only, with reference to the accompanying drawings, in which:

Figure 1 shows a schematic of the treatment of a genomic DNA clone with a frequent cutting restriction endonuclease, such as Sau3A1, followed by ligation of adaptors to restriction fragments bearing specific primer sequences - all fragments are dealt with simultaneously, but for simplicity only one is shown;
Figure 2 shows a schematic of an amplification step, following the steps of Figure 1, in which fragments are amplified by PCR using adaptor primers;
Figure 3 shows a step following the step of Figure 2, in which amplified fragments are subdivided into 10, wells, each well being identified by a pair of primers used to sort added molecules, each well initially containing one of the pair of primers, there being 4 primers

each with one base probe sequence and each well having 1 of 10 possible pairs generated by a combination of the four primers, the second primer being added after one cycle of synthesis of the first;

Figure 4 shows a schematic of a differential amplification step, following the step of Figure 3, in which the contents of a well containing a primer terminated with AC followed by a probe terminated by AG is amplified and then one cycle of synthesis is performed with the first primer and double strands captured with avidinated beads;

Figure 5 shows a schematic of steps subsequent to those of Figure 4, in which the non-immobilised strand is melted off and washed away and the reaction residue polymerised, a second primer then being added and a second cycle of synthesis performed; and

Figures 6A and 6B show a schematic of steps subsequent to those of Figure 5, in which the non-immobilised strand is melted off and transferred to a fresh reaction vessel, and both primers are then added to the fresh free strand to amplify by PCR.

In the present invention, the nucleic acid population is not isolated (such as by capture onto a solid phase) prior to contacting it with the oligonucleotide sequence(s). Thus each nucleic acid in the population may initially move freely in the suspension or solution in which it is contained. After contacting the nucleic acid population with the oligonucleotide sequence(s), preferably only the nucleic acid(s) which have correctly hybridised to the oligonucleotide sequence(s) are isolated (preferably by capture onto a solid phase).

In more detail, the method of this invention may comprise the following steps:

1. Restricting a large nucleic acid or population of large nucleic acids to generate fragments with known termini.

2. Ligating adaptors or linkers to the termini of these nucleic acid molecules. The ligated adaptor provides a known sequence at the termini of a population of nucleic acids which can be used to design primers which extend beyond the terminal adaptor sequence into unknown sequence adjacent to the known adaptor sequence allowing the unknown sequence to be probed.

3. Optionally amplifying the adaptored fragments using primers complementary to the whole or part of the adaptor sequences at the termini of the adaptored fragments.
4. Optionally normalising the population of adaptored nucleic acids.
5. Selectively amplifying subsets of the nucleic acids through the use of pairs of primers which partially overlap into the unknown sequence. The overlapping primer will hybridise to a subset of the whole population. The size of the subset is determined by the length of overlap of the primer into the adjacent sequence.

The methods of this invention may be applied cyclically to sub-populations of sorted nucleic acids generated by the methods of this invention. Each cycle further reduces the complexity of the population. If necessary the cycles can be repeated until unique nucleic acid is obtained.

In a preferred embodiment the step of restricting nucleic acid is coupled to the ligation of adapters. Preferred restriction endonucleases for use with this invention cleave within their recognition sequence generating sticky-ends that do not encompass the whole recognition sequence. This allows adapters to be designed that bear sticky ends complementary to those generated by the preferred restriction endonuclease but which do not regenerate the recognition site of the preferred restriction endonuclease. This means that if the restriction reaction is performed in the presence of ligase and adapters, the ligation of restriction fragments to each other is reduced by continuous cleavage of these ligations whereas ligation of adapters is irreversible so the presence of adapters drives the restriction to completion and similarly the restriction endonuclease drives the ligation reaction to completion. This process ensures that a very high proportion of restriction fragments are ligated to adaptors. This is advantageous as ligation of adapters to restriction fragments is a relatively inefficient process. This is due to random ligation of restriction products to each other if these are phosphorylated. In this embodiment the adapters used are preferably not phosphorylated at their 5' hydroxyl groups so that they cannot ligate to themselves.

GB 9115407.0 describes a method of normalising a population of nucleic acids comprising the following steps:

1. Combining a mixture of heterogeneous DNA fragments with oligonucleotide primers compatible with some nucleic acid amplification system and denaturing the double stranded heterogeneous DNA.

2. Altering the conditions, i.e. reducing the temperature, to allow the more common species to re-anneal while preventing the primers from annealing to the DNA. The temperature for re-annealing at this stage must be higher than the melting temperature of the PCR primers.

3. Altering the reaction conditions further to allow the PCR primers to anneal to the remaining single stranded DNA which should represent the rarer species.

4. Performing strand extension of the primed species.

Advantageously, the above steps are applied cyclically a number of times to amplify the rarer species to a significant extent.

Application of this method to sequences with known termini permits the design of primers with very specific melting temperatures allowing the method to be used generically. Use of this method is particularly advantageous in reducing the complexity of genomic DNA as a significant proportion of most genomic DNA is repetitive sequence.

The advantage of providing a known sequence adjacent to probe sequence allows one to design libraries of probes, where all the probes in a library have the same melting temperature. This is advantageous as hybridisation of the entire library can be performed simultaneously at a single temperature whilst retaining the stringency of hybridisation.

Consider a large DNA fragment such as a mitochondrial genome or a cosmid or a microbial genome. To perform steps 1 to 4 of the method described above, such a large molecule can be cleaved with a frequently cutting restriction enzyme to generate fragments of the order of a few hundred bases in length. If a restriction endonuclease like Sau3A1 is used fragments with a

known sticky end are left, to which double stranded adaptors can be ligated. These adaptors will bear a known primer sequence, and a sticky end complementary to that produced by the restriction endonuclease to permit ligation. A combined restriction and ligation protocol as described above is appropriate.

The majority of properly restricted fragments as a result bear an adaptor at each of their termini. This permits amplification of the adaptored restriction fragments at this stage if that is desired. After adaptoring and any non-selective amplification and normalisation, the nucleic acids can be differentially amplified to generate specific subsets of the starting population. The method of differential amplification preferably comprises the following steps:

1. Dividing the adaptored population of restriction fragments into separate wells. If, for example, primers with an overlap of a single base are used then the adaptored fragments would be divided into 10 or 16 wells.
2. Adding to each well one type of biotinylated primer of a predetermined set. The primer bears a sequence complementary to that provided by the adaptor and restriction site. The primer additionally bears an overlap of a predetermined number of bases beyond the known sequence into the unknown sequence immediately adjacent to the restriction site. Primers with different overlaps are added to different well. Four primers are need if a 1 base overlap is used. If 16 wells are used each of the 4 primers are added to 4 wells.
3. Denaturing the amplified fragment population that was subdivided into each well by raising the temperature. The temperature is then reduced to permit the primer sequences to anneal. Primers preferably have equalised melting temperatures so that conditions for use of all primers are the same.
4. Adding thermostable polymerase and nucleotides to extend annealed primers.
5. Capturing the biotinylated strand extension products from (4) onto a solid phase substrate derivitised with avidin. This may be effected through the addition of avidinated beads. These may optionally be magnetic beads.
6. Melting off the non-biotinylated complementary strand and washing this away. This leaves a single stranded copy of the selected fragments immobilised on the solid phase support.

7. To each of the separate pools is added one of the same set of primers as used in step (2) but not biotinylated, such that each pool receives a different combination of primers from this step and step (2). The primers should anneal to the single stranded capture molecules from (6). If 16 pools are used, to each is added one of the same 4 primers, but not biotinylated such that each of the 16 pools carries one of the possible different combinations of pairs of the 4 primers.
8. Extending the primed captured strands with polymerase and nucleotide triphosphates.
9. Denaturing the free strand from the captured strand by raising the temperature. The 'selected' free strand is thus released into solution. The liquid phase can be transferred to fresh reaction vessel or the solid phase support bearing the captured strands from (5) can be removed. This is very easy if the support used are magnetic beads as these can be removed by electromagnetic attraction to a probe.

The isolated free strands from (9) are thus isolated. At this stage the selected strands can be captured onto a solid phase support or amplified or the process of differential amplification can be repeated on the isolated subsets generated to further sub-sort these populations. This would be effected by using primers which overlap further into the unknown sequence adjacent to the known sequence of the adapter and the selected fragment. The sorted fragment could equally be cloned into a biological vector at this stage if desired.

Generating a captured library is advantageous in that it facilitates easy manipulation of the library of fragments. Such manipulations include copying, amplification and probing of the library for particular sequences. A captured library dispenses with any requirement for biological cloning vectors to maintain the library as such a library can be readily copied using polymerases and nucleotide triphosphates. The captured library can be readily washed and can very easily be stored in a refrigerated environment.

It should be noted in the example of primers that overlap by a single base, that the amplification products from the well containing a primer terminated by A followed by the primer terminated by G gives the complement of the well where G is followed by A. It might therefore be

desirable to pool the reactions of where the same pair of primers are present but used in a different order to ensure that both strands of each DNA molecule are present and captured on the solid phase support. This would thus give 10 different pools. This is a convenient number as one can reduce the complexity of a library by one order of magnitude with four primers. Each sorted library of fragments can be further sub-sorted to an arbitrary degree.

An alternative embodiment of this method uses primers already immobilised on a solid phase support, preferably covalently linked to the support instead of biotinylated primers in step (2) of the differential amplification process. Such solid phase supports can be magnetic beads, as described in EP-A-0 091 453 and EP-A-0 106 873, or the support could be polymer beads. PCT/GB92/02394 describes a solid phase polymer support in a micro-column where the solid phase support are beads of silica gel. The beads are retained between two frits in the column through which solvents and reagents can flow. Such apparatus is also applicable with this invention.

One can clearly repeat the sorting process starting from a captured library that has been previously sorted.

One can also clearly use just 10 wells to generate sorted populations as all of the sequence information in a series of 16 wells will be present if just the 10 different pairs of primer combinations are used.

It should also be clear that labels can be introduced into sorted molecules by the primers used as part of the sorting process. Methods of introducing labels into primer oligonucleotides are well known in the art. Biotin has been discussed above, but many others are applicable.

One can also use probes which overlap beyond the provided adaptor sequence to any extent. It becomes more difficult, however, to ensure the stringency of hybridisation as the number of bases extending into the unknown sequence from the adaptor is increased.

To effect higher degrees of sorting one can either sort a sorted library with a set of four primers that overlap beyond the known terminal sequences by a single base or one can use primers with a longer sequence overlap. To sort an adaptored population of nucleic acid fragments using primers with a 2 base overlap beyond the adaptor sequence, the adaptored population of restriction fragments is sub-divided into 256 wells. In each well is one of 16 biotinylated primers which bear a sequence complementary to that provided by the adaptor and restriction site. The primers additionally bear an overlap of 2 bases beyond the known sequence into the unknown sequence immediately adjacent to the restriction site. The amplified fragment population subdivided into each well is denatured by raising the temperature and cooled allowing the primer sequences to anneal. Primers, again, preferably have equalised melting temperatures so that conditions for use of all primers is the same. Thermostable polymerase and nucleotides are added to extend annealed primers. Biotinylated fragments are captured onto a solid phase substrate via avidin and the complementary strand is melted off and washed away. To each of the 256 pools is added one of the same 16 primers, but not biotinylated such that each of the 256 pools carries one of the possible different combinations of pairs of the 16 primers. Again, AT followed by GC gives the complement of the reaction of GC followed by AT so it might be desirable to pool these pairs to give a total of 136 pools. For an overlap of n bases, one can distinguish 4^n distinct sequences. If both termini of a molecule are used to select fragments then one can distinguish fragments into $(4^n \times (4^n+1)/2)$ distinct sets, since the orientation of each fragment is unknown.

Sorting a library resolves fragments from a large, complex population into defined sets whose size will be statistically regular and determinable as long as the size of the parent library is known, even if only approximately. The composition of the sorted library will be less complex than that of the parent library. This allows for useful manipulations of a large library without loss of information as all the sequences present in the starting library should be present in one of the sub libraries as long as long as all of the possible sub-libraries are generated. This

method offers greater ease of manipulation of complex nucleic acid libraries and greater precision of manipulation than cloning into biological vectors.

To put this invention into practise requires the construction of probe oligonucleotides (ONs). Precise control over hybridisation conditions will be required to ensure clean results in differential amplification.

Details and reviews on the construction of labelled and modified ONs are available in numerous up-to-date texts, see references 1 to 6 below. A brief discussion of preferred design possibilities is given below.

There are major differences between the stability of short oligonucleotide duplexes containing all Watson-Crick base pairs. For example, duplexes comprising only adenine and thymine are unstable relative to duplexes of guanine and cytosine only. These differences in stability can present problems when trying to hybridise mixtures of short oligonucleotides to a target RNA. Low temperatures are needed to hybridise A-T rich sequences but at these temperatures G-C rich sequences will hybridise to sequences that are not fully complementary. This means that some mismatches may occur, and specificity can be lost for the G-C rich sequences. At higher temperatures G-C rich sequences will hybridise specifically but A-T rich sequences will not hybridise.

It is desirable that probes within a library behave in a similar manner, i.e. they should have similar melting temperatures and preferably also binding kinetics. In order to normalise these effects, modifications can be made to nucleic acids. Modifications fall into three broad categories: base modifications, backbone modifications and sugar modifications.

Base modifications

Numerous modifications can be made to the standard Watson-Crick bases. The following are examples of modifications that should normalise base pairing energies to some extent but they are not limiting:

•The adenine analogue 2,6-diaminopurine forms three hydrogen bonds to thymine rather than two and therefore forms more stable base pairs.

•The thymine analogue 5-propynyl dU forms more stable base pairs with adenine.

•The guanine analogue hypoxanthine forms two hydrogen bonds with cytosine rather than three and therefore forms less stable base pairs.

These and other possible modifications should make it possible to compress the temperature range at which short oligonucleotides can hybridise specifically to their complementary sequences.

Backbone modifications

Nucleotides may be readily modified in the phosphate moiety. Under certain conditions, such as low salt concentration, analogues such as methylphosphonates, triesters and phosphoramidates have been shown to increase duplex stability. Such modifications may also have increased nuclease resistance. Further phosphate modifications include phophodithirates and boranophosphates, each of which increase the stability of ONs.

Isosteric replacement of phosphorus by sulphur gives nuclease resistant ONs (reference 7). Replacement by carbon at either phosphorus or linking oxygen is also a further possibility.

Sugar modifications

Various modifications to the 2' position in the sugar moiety may be made (references 12 and 13). The sugar may be replaced by a different sugar such as hexose or the entire sugar phosphate backbone can be entirely replaced by a novel structure such as in peptide nucleic acids (PNA). For a discussion see reference 8. PNA may be the ideal choice as it forms duplexes of the highest thermal stability of any analogues so far discovered.

Artificial mismatches

One major source of error in hybridisation reactions is the stringency of hybridisation of the primers to the target sequence and to the unknown bases beyond. If the primers designed for a target bear single artificially introduced mismatches the discrimination of the system is much higher (Zhen Guo *et al.*, Nature Biotechnology 15. 331-335, April 1997). Additional mismatches are not tolerated to the same extent that a single mismatch would be when a fully complementary primer is used. Thus this can be exploited in the method disclosed above. If the probe used to extends beyond the provided sequence by 1 base, an artificial mismatch, 1 helical turn away from the probe base destabilises the double helix to a considerable degree if there is a second mismatch at the probe site.

Details on effects of hybridisation conditions for nucleic acid probes can be found in references 9 to 11.

Mass labels for use in the present invention are discslosed in patent application PCT/GB98/00127. Further labels for use in the present invention are discussed in the UK applications of Page White & Farrer file numbers 87820, 87821, 87900.

References

(1) Gait, M.J. editor, 'Oligonucleotide Synthesis: A Practical Approach', IRL Press. Oxford. 1990

(2) Eckstein. editor, 'Oligonucleotides and Analogues: A Practical Approach'. IRL Press. Oxford. 1991

(3) Kricka. editor. 'Nonisotropic DNA Probe Techniques'. Academic Press, San Diego, 1992

(4) Haugland, 'Handbook of Fluorescent Probes and Research Chemicals', Molecular Probes, Inc., Eugene. 1992

(5) Keller and Manack, 'DNA Probes, 2nd Edition', Stockton Press, New York, 1993

(6) Kessler, editor, 'Nonradioactive Labelling and Detection of Biomolecules', Springer-Verlag, Berlin, 1992.

(7) J.F. Milligan, M.D. Matteucci, J.C. Martin, J. Med. Chem. 36(14), 1923 - 1937,1993.

(8) P.E. Nielsen, Annu. Rev. Biophys. Biomol. Struct. 24, 167 - 183, 1995.

(9) Wetmur. Critical Reviews in Biochemistry and Molecular Biology, 26, 227-259, 1991

(10) Sambrook et al, 'Molecular Cloning: A Laboratory Manual, 2nd Edition', Cold Spring Harbour Laboratory, New York, 1989

(11) Hames, B.D., Higgins, S.J., 'Nucleic Acid Hybridisation: A Practical Approach', IRL Press, Oxford, 1988

(12) C.J. Guinosso, G.D. Hoke, S.M. Freier, J.F. Martin, D.J. Ecker, C.K. Mirabelle, S.T. Crooke, P.D. Cook, Nucleosides Nucleotides 10, 259 - 262, 1991.

(13) M. Carmo-Fonseca, R. Pepperkok, B.S. Sproat, W. Ansorge, M.S. Swanson, A.I. Lamond, EMBO J. 7, 1863 - 1873, 1991.

Claims:

1. A method for categorising nucleic acid, which method comprises producing a nucleic acid population by action of an endonuclease on double-stranded nucleic acid, such that each nucleic acid in the nucleic acid population has a double-stranded portion, contacting the nucleic acid population with one or more oligonucleotide sequences, and isolating nucleic acid which correctly hybridises to an oligonucleotide sequence by capturing the oligonucleotide sequence on a solid phase, wherein each oligonucleotide sequence has a pre-determined recognition sequence, the nucleic acid being categorised by its ability to correctly hybridise to oligonucleotide sequences having the recognition sequence, the recognition sequence being situated such that it recognises a sequence in the double-stranded portion of the nucleic acid, one or more different recognition sequences being represented in the oligonucleotide sequences.

2. A method according to claim 1, wherein the endonuclease is selected such that each nucleic acid in the nucleic acid population has a sticky end of a known common length extending from a terminal of its double-stranded portion.

3. A method according to claim 1, wherein the endonuclease is selected such that each sticky end of each nucleic acid in the nucleic acid population has the same known base sequence.

4. A method according to claim 3, wherein prior to contacting the nucleic acid population with the oligonucleotide sequences, the nucleic acid population is contacted with an adaptor to ligate the adaptor to a terminal of each nucleic acid in the nucleic acid population, wherein the adaptor comprises a double-stranded primer portion having a known base sequence, and a single-stranded portion complementary to the known sticky end of the nucleic acids in the nucleic acid population.

5. A method according to claim 4, wherein each oligonucleotide sequence comprises a first sequence, a second sequence attached to the first sequence and a third sequence attached to the second sequence, in which the first sequence is complementary to the sequence of the primer portion of the adaptor, the second sequence is complementary to the known sticky end

of the nucleic acids in the nucleic acid population, and the third sequence comprises the pre-determined recognition sequence.

6. A method according to claim 2, wherein the endonuclease is selected such that the sticky ends of the nucleic acids in the nucleic acid population have a plurality of different base sequences.

7. A method according to claim 6, wherein prior to contacting the nucleic acid population with the oligonucleotide sequences, the nucleic acid population is contacted with an array of adaptors to ligate an adaptor to a terminal of the nucleic acids in the nucleic acid population, wherein each adaptor comprises a double-stranded primer portion having a known base sequence, and a single-stranded portion of the same length as the sticky ends of the nucleic acids in the nucleic acid population, all of the possible base sequences of the single-stranded portion of the adaptor being represented in the array of adaptors.

8. A method according to claim 7, wherein each oligonucleotide sequence comprises a first sequence, a second sequence attached to the first sequence and a third sequence attached to the second sequence, in which the first sequence is complementary to the sequence of the primer portion of the adaptors, the second sequence is of the same length as the sticky ends of the nucleic acids in the nucleic acid population, and the third sequence comprises the pre-determined recognition sequence, and wherein in any one group of oligonucleotides having the same recognition sequence all of the possible base sequences of the second sequence are represented.

9. A method according to claim 5 or claim 8, wherein the recognition sequence consists of one base.

10. A method according to claim 5 or claim 8, wherein the recognition sequence consists of two or more bases.

11. A method according to any of claims 5 and 8-10, wherein in any one group of oligonucleotides having the same recognition sequence the third sequence consists of the

recognition sequence and a pre-determined number of bases situated between the second sequence and the recognition sequence, all possible sequences of the pre-determined number of bases in the third sequence being represented in that group of oligonucleotides.

12. A method according to any preceding claim, wherein the nucleic acid population is amplified by PCR prior to reaction with the oligonucleotide sequences.

13. A method according to any preceding claim, wherein those nucleic acids are isolated both terminals of which correctly hybridise to an oligonucleotide sequence.

14. A method according to claim 13, wherein a first set of oligonucleotide sequences is contacted with the nucleic acid population in a first step by denaturing the nucleic acid population in the presence of the first set of sequences to produce a single-stranded nucleic acid population and allowing the single-stranded nucleic acid to hybridise to the first sequences, immobilising those nucleic acids which correctly hybridise to the first sequences, extending the correctly hybridised oligonucleotide sequences along the single-stranded portion of the immobilised nucleic acid to form double-stranded nucleic acid, denaturing the double-stranded nucleic acid and removing non-immobilised species to isolate the resulting immobilised single-stranded nucleic acid, contacting the immobilised single-stranded nucleic acid with a second set of oligonucleotide sequences in a second step, extending the correctly hybridised oligonucleotide sequences along the immobilised single-stranded nucleic acid to form double-stranded nucleic acid, denaturing the double-stranded nucleic acid and isolating the resulting non-immobilised single-stranded nucleic acid.

15. A method according to claim 14, wherein the extended and isolated products of the first step and/or the extended and isolated products of the second step are amplified by PCR.

16. A method according to claim 14 or claim 15, wherein the correctly hybridised nucleic acids are immobilised by immobilising the oligonucleotide sequences.

AMENDED SHEET

17. A method according to claim 16, wherein each oligonucleotide in the first set of sequences carries a biotin residue such that prior to or after hybridising to the nucleic acid the sequence is captured on an avidinated solid phase.

18. A method according to claim 16, wherein each oligonucleotide in the first set of sequences is covalently attached to a solid support prior to contacting with the nucleic acid population.

19. A method according to any of claims 14-18, wherein the recognition sequence of the first and second set of oligonucleotide sequences consists of one base and, prior to performing the first step, the nucleic acid population is sub-divided into 16 wells, each well containing oligonucleotides from the first set of sequences having one of the four possible recognition sequences, and wherein in the second step oligonucleotides from the second set of sequences are added to each well, such that all possible combinations of the identities of the first and second set of oligonucleotide sequences and their order of addition to the well are represented in the 16 wells.

20. A method according to any of claims 14-18, wherein the recognition sequence of the first and second set of oligonucleotide sequences consists of two bases and, prior to performing the first step, the nucleic acid population is sub-divided into 256 wells, each well containing oligonucleotides from the first set of sequences having one of the 16 possible recognition sequences, and wherein in the second reaction oligonucleotides from the second set of sequences are added to each well, such that all possible combinations of the identity of the first and second set of oligonucleotide sequences and their order of addition to the well are represented in the 256 wells.

21. A method according to claim 19, wherein the contents of each pair of wells to which the same pair of oligonucleotide sequences were added but in a different order, are combined to give 10 different wells.

22. A method according to claim 20, wherein the contents of each pair of wells to which the same pair of oligonucleotide sequences were added but in a different order, are combined to give 136 different wells.

AMENDED SHEET

23. A method according to any preceding claim, wherein the oligonucleotide sequences have equalised melting temperatures.

24. A method according to claim 23, wherein the melting temperatures are equalised by incorporating one or more analogues of natural nucleotides into the oligonucleotide sequences, the analogues comprising base modifications, sugar modifications and/or backbone modifications.

25. A method according to any preceding claim, wherein the endonuclease is selected such that it cuts the nucleic acid at a site within the recognition site of the endonuclease.

26. A kit for categorising a nucleic acid, comprising one or more adaptors and one or more sets of oligonucleotide sequences, wherein the adaptors comprise nucleic acid having a double-stranded primer portion of a known sequence and a single-stranded portion of a pre-determined length, either each single-stranded portion of each nucleic acid in the adaptors having the same pre-determined sequence or all possible sequences of the single-stranded portion being represented in the adaptors, and wherein each oligonucleotide sequence comprises a first sequence, a second sequence attached to the first sequence and a third sequence attached to the second sequence, in which the first sequence is complementary to the sequence of the primer portion of the adaptor, the second sequence is the same sequence as the single-stranded portion of the adaptors or all possible second sequences of the same length as the single-stranded portion of the adaptors are represented within the set of oligonucleotides, and the third sequence comprises a pre-determined recognition sequence.

27. A kit according to claim 26, wherein the recognition sequence consists of one base.

28. A kit according to claim 26, wherein in the recognition sequence consists of two or more bases.

29. A kit according to any of claims 26-28, wherein in any one group of oligonucleotides having the same recognition sequence, the third sequence consists of the recognition sequence and a pre-

AMENDED SHEET

determined number of bases situated between the second sequence and the recognition sequence, all of the possible sequences of the pre-determined number of bases in the third sequence being represented in that group of oligonucleotides.

30. A kit according to any of claims 26-29, comprising two sets of oligonucleotide sequences, each of the oligonucleotides in one set being biotinylated.

31. A kit according to any of claims 26-29, comprising two sets of oligonucleotide sequences, each of the oligonucleotides in one set being covalently attached to a solid support.

32. A kit according to any of claims 26-31, additionally comprising an endonuclease.

33. A kit according to claim 32, wherein the endonuclease is selected such that when it is reacted with double-stranded nucleic acid, nucleic acids are produced each of which comprises a double-stranded portion.

34. A kit according to claim 33, wherein the endonuclease is selected such that the nucleic acids produced have a sticky end of a known common length extending from a terminal of the double-stranded portion, and wherein each sticky end of each nucleic acid in the nucleic acid population has the same known base sequence.

35. A kit according to claim 33, wherein the endonuclease is selected such that the nucleic acids produced have a sticky end of a known common length extending from a terminal of the double-stranded portion, and wherein the sticky ends of the nucleic acids in the nucleic acid population exhibit a plurality of different base sequences.

36. A kit according to any of claims 26-35, wherein the endonuclease is selected such that it cuts the nucleic acid at a site within the recognition site of the endonuclease.

AMENDED SHEET.

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

PCT

NOTIFICATION OF ELECTION

(PCT Rule 61.2)

To:

United States Patent and Trademark Office
(Box PCT)
Crystal Plaza 2
Washington, DC 20231
ÉTATS-UNIS D'AMÉRIQUE

in its capacity as elected Office

Date of mailing (day/month/year) 08 April 1999 (08.04.99)	
International application No. PCT/GB98/02043	**Applicant's or agent's file reference** 86911/JND/CH
International filing date (day/month/year) 13 July 1998 (13.07.98)	**Priority date** (day/month/year) 11 July 1997 (11.07.97)
Applicant SCHMIDT, Gunter et al	

1. The designated Office is hereby notified of its election made:

[X] in the demand filed with the International Preliminary Examining Authority on:

05 February 1999 (05.02.99)

[ ] in a notice effecting later election filed with the International Bureau on:

2. The election [X] was
[ ] was not

made before the expiration of 19 months from the priority date or, where Rule 32 applies, within the time limit under Rule 32.2(b).

The International Bureau of WIPO **34, chemin des Colombettes** **1211 Geneva 20, Switzerland** Facsimile No.: (41-22) 740.14.35	Authorized officer Lazar Joseph Panakal Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

PCT

INTERNATIONAL SEARCH REPORT

(PCT Article 18 and Rules 43 and 44)

Applicant's or agent's file reference	FOR FURTHER ACTION	see Notification of Transmittal of International Search Report (Form PCT/ISA/220) as well as, where applicable, item 5 below.
86911/JND/CH		
International application No.	International filing date (*day/month/year*)	(Earliest) Priority Date (*day/month/year*)
PCT/GB 98/ 02043	13/07/1998	11/07/1997
Applicant		
BRAX GENOMICS LIMITED et al.		

This International Search Report has been prepared by this International Searching Authority and is transmitted to the applicant according to Article 18. A copy is being transmitted to the International Bureau.

This International Search Report consists of a total of 3 sheets.

[X] It is also accompanied by a copy of each prior art document cited in this report.

1. [ ] **Certain claims were found unsearchable** (see Box I).

2. [ ] **Unity of invention is lacking** (see Box II).

3. [ ] The international application contains disclosure of a **nucleotide and/or amino acid sequence listing** and the international search was carried out on the basis of the sequence listing
   - [ ] filed with the international application.
   - [ ] furnished by the applicant separately from the international application,
     - [ ] but not accompanied by a statement to the effect that it did not include matter going beyond the disclosure in the international application as filed.
   - [ ] Transcribed by this Authority

4. With regard to the **title**,
   - [X] the text is approved as submitted by the applicant
   - [ ] the text has been established by this Authority to read as follows:

5. With regard to the **abstract**,
   - [X] the text is approved as submitted by the applicant
   - [ ] the text has been established, according to Rule 38.2(b), by this Authority as it appears in Box III. The applicant may, within one month from the date of mailing of this International Search Report, submit comments to this Authority.

6. The figure of the **drawings** to be published with the abstract is:

   Figure No. 1,2
   - [ ] as suggested by the applicant.
   - [X] because the applicant failed to suggest a figure.
   - [ ] because this figure better characterizes the invention.

   [ ] None of the figures.

Form PCT/ISA/210 (first sheet) (July 1992)

INTERNATIONAL SEARCH REPORT

International Application No
PCT/GB 98/02043

A. CLASSIFICATION OF SUBJECT MATTER
IPC 6 C12Q1/68

According to International Patent Classification (IPC) or to both national classification and IPC

B. FIELDS SEARCHED

Minimum documentation searched (classification system followed by classification symbols)
IPC 6 C12Q

Documentation searched other than minimum documentation to the extent that such documents are included in the fields searched

Electronic data base consulted during the international search (name of data base and, where practical, search terms used)

C. DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication, where appropriate, of the relevant passages	Relevant to claim No.
X	WONG D M ET AL: "BRANCH CAPTURE REACTIONS: DISPLACERS DERIVED FROM ASYMMETRIC PCR" NUCLEIC ACIDS RESEARCH, vol. 19, no. 9, 11 May 1991, pages 2251-2259, XP000204316 see whole document and esp. figure 1 ---	1
X	GUILFOYLE R. A. ET AL.,: "Ligation-mediated PCR amplification of specific fragments from a class-II restriction endonuclease total digest" NUCLEIC ACIDS RESEARCH, vol. 25, no. 9, - 1 May 1997 pages 1854-1858, XP002076198	1-13, 25-36
Y	see the whole document ---	14-24
	-/--	

[X] Further documents are listed in the continuation of box C.

[X] Patent family members are listed in annex.

° Special categories of cited documents :

"A" document defining the general state of the art which is not considered to be of particular relevance

"E" earlier document but published on or after the international filing date

"L" document which may throw doubts on priority claim(s) or which is cited to establish the publication date of another citation or other special reason (as specified)

"O" document referring to an oral disclosure, use, exhibition or other means

"P" document published prior to the international filing date but later than the priority date claimed

"T" later document published after the international filing date or priority date and not in conflict with the application but cited to understand the principle or theory underlying the invention

"X" document of particular relevance; the claimed invention cannot be considered novel or cannot be considered to involve an inventive step when the document is taken alone

"Y" document of particular relevance; the claimed invention cannot be considered to involve an inventive step when the document is combined with one or more other such documents, such combination being obvious to a person skilled in the art.

"&" document member of the same patent family

Date of the actual completion of the international search	Date of mailing of the international search report
20 October 1998	29/10/1998

Name and mailing address of the ISA	Authorized officer
European Patent Office, P.B. 5818 Patentlaan 2 NL - 2280 HV Rijswijk Tel. (+31-70) 340-2040, Tx. 31 651 epo nl, Fax: (+31-70) 340-3016	Müller, F

1

Form PCT/ISA/210 (second sheet) (July 1992)

C.(Continuation) DOCUMENTS CONSIDERED TO BE RELEVANT

Category °	Citation of document, with indication,where appropriate, of the relevant passages	Relevant to claim No.
Y	EP 0 370 694 A (EASTMAN KODAK CO ;CETUS CORP (US)) 30 May 1990 see esp. claim 2	14-24
A	EP 0 735 144 A (JAPAN RES DEV CORP) 2 October 1996 see the whole document	1-36
A	WO 94 01582 A (MEDICAL RES COUNCIL ;SIBSON DAVID ROSS (GB)) 20 January 1994 cited in the application see the whole document	1-36
A	US 5 508 169 A (DEUGAU KENNETH V ET AL) 16 April 1996 cited in the application see the whole document	1-36

1

INTERNATIONAL SEARCH REPORT

Information on patent family members

International Application No
PCT/GB 98/02043

Patent document cited in search report	Publication date	Patent family member(s)	Publication date
EP 0370694 A	30-05-1990	CA 2002076 A	21-05-1990
		DK 582189 A	22-05-1990
		JP 2299600 A	11-12-1990
EP 0735144 A	02-10-1996	JP 2763277 B	11-06-1998
		JP 9028399 A	04-02-1997
		JP 2763278 B	11-06-1998
		JP 8322598 A	10-12-1996
		AU 692685 B	11-06-1998
		AU 5031196 A	10-10-1996
		US 5707807 A	13-01-1998
WO 9401582 A	20-01-1994	AT 159986 T	15-11-1997
		AU 686563 B	12-02-1998
		AU 4575893 A	31-01-1994
		CA 2139944 A	20-01-1994
		DE 69315074 D	11-12-1997
		DE 69315074 T	05-03-1998
		EP 0650528 A	03-05-1995
		JP 7508883 T	05-10-1995
		US 5728524 A	17-03-1998
US 5508169 A	16-04-1996	CA 2036946 A	07-10-1991

Claims:

1. A method for categorising nucleic acid, which method comprises producing a nucleic acid population by action of an endonuclease on double-stranded nucleic acid, such that each nucleic acid in the nucleic acid population has a double-stranded portion, contacting the nucleic acid population with one or more oligonucleotide sequences, and isolating nucleic acid which correctly hybridises to an oligonucleotide sequence, wherein each oligonucleotide sequence has a pre-determined recognition sequence, the nucleic acid being categorised by its ability to correctly hybridise to oligonucleotide sequences having the recognition sequence, the recognition sequence being situated such that it recognises a sequence in the double-stranded portion of the nucleic acid, one or more different recognition sequences being represented in the oligonucleotide sequences.

2. A method according to claim 1, wherein the endonuclease is selected such that each nucleic acid in the nucleic acid population has a sticky end of a known common length extending from a terminal of its double-stranded portion.

3. A method according to claim 1, wherein the endonuclease is selected such that each sticky end of each nucleic acid in the nucleic acid population has the same known base sequence.

4. A method according to claim 3, wherein prior to contacting the nucleic acid population with the oligonucleotide sequences, the nucleic acid population is contacted with an adaptor to ligate the adaptor to a terminal of each nucleic acid in the nucleic acid population, wherein the adaptor comprises a double-stranded primer portion having a known base sequence, and a single-stranded portion complementary to the known sticky end of the nucleic acids in the nucleic acid population.

5. A method according to claim 4, wherein each oligonucleotide sequence comprises a first sequence, a second sequence attached to the first sequence and a third sequence attached to the second sequence, in which the first sequence is complementary to the sequence of the primer portion of the adaptor, the second sequence is complementary to the known sticky end of the nucleic acids in the nucleic acid population, and the third sequence comprises the pre-determined recognition sequence.

6. A method according to claim 2, wherein the endonuclease is selected such that the sticky ends of the nucleic acids in the nucleic acid population have a plurality of different base sequences.

7. A method according to claim 6, wherein prior to contacting the nucleic acid population with the oligonucleotide sequences, the nucleic acid population is contacted with an array of adaptors to ligate an adaptor to a terminal of the nucleic acids in the nucleic acid population, wherein each adaptor comprises a double-stranded primer portion having a known base sequence, and a single-stranded portion of the same length as the sticky ends of the nucleic acids in the nucleic acid population, all of the possible base sequences of the single-stranded portion of the adaptor being represented in the array of adaptors.

8. A method according to claim 7, wherein each oligonucleotide sequence comprises a first sequence, a second sequence attached to the first sequence and a third sequence attached to the second sequence, in which the first sequence is complementary to the sequence of the primer portion of the adaptors, the second sequence is of the same length as the sticky ends of the nucleic acids in the nucleic acid population, and the third sequence comprises the pre-determined recognition sequence, and wherein in any one group of oligonucleotides having the same recognition sequence all of the possible base sequences of the second sequence are represented.

9. A method according to claim 5 or claim 8, wherein the recognition sequence consists of one base.

10. A method according to claim 5 or claim 8, wherein the recognition sequence consists of two or more bases.

11. A method according to any of claims 5 and 8-10, wherein in any one group of oligonucleotides having the same recognition sequence the third sequence consists of the recognition sequence and a pre-determined number of bases situated between the second sequence and the recognition sequence, all possible sequences of the pre-determined number of bases in the third sequence being represented in that group of oligonucleotides.

12. A method according to any preceding claim, wherein the nucleic acid population is amplified by PCR prior to reaction with the oligonucleotide sequences.

13. A method according to any preceding claim, wherein those nucleic acids are isolated both terminals of which correctly hybridise to an oligonucleotide sequence.

14. A method according to claim 13, wherein a first set of oligonucleotide sequences is contacted with the nucleic acid population in a first step by denaturing the nucleic acid population in the presence of the first set of sequences to produce a single-stranded nucleic acid population and allowing the single-stranded nucleic acid to hybridise to the first sequences, immobilising those nucleic acids which correctly hybridise to the first sequences, extending the correctly hybridised oligonucleotide sequences along the single-stranded portion of the immobilised nucleic acid to form double-stranded nucleic acid, denaturing the double-stranded nucleic acid and removing non-immobilised species to isolate the resulting immobilised single-stranded nucleic acid, contacting the immobilised single-stranded nucleic acid with a second set of oligonucleotide sequences in a second step, extending the correctly hybridised oligonucleotide sequences along the immobilised single-stranded nucleic acid to

form double-stranded nucleic acid, denaturing the double-stranded nucleic acid and isolating the resulting non-immobilised single-stranded nucleic acid.

15. A method according to claim 14, wherein the extended and isolated products of the first step and/or the extended and isolated products of the second step are amplified by PCR.

16. A method according to claim 15 or claim 16, wherein the correctly hybridised nucleic acids are immobilised by immobilising the oligonucleotide sequences.

17. A method according to claim 16, wherein each oligonucleotide in the first set of sequences carries a biotin residue such that prior to or after hybridising to the nucleic acid the sequence is captured on an avidinated solid phase.

18. A method according to claim 16, wherein each oligonucleotide in the first set of sequences is covalently attached to a solid support prior to contacting with the nucleic acid population.

19. A method according to any of claims 14-18, wherein the recognition sequence of the first and second set of oligonucleotide sequences consists of one base and, prior to performing the first step, the nucleic acid population is sub-divided into 16 wells, each well containing oligonucleotides from the first set of sequences having one of the four possible recognition sequences, and wherein in the second step oligonucleotides from the second set of sequences are added to each well, such that all possible combinations of the identities of the first and second set of oligonucleotide sequences and their order of addition to the well are represented in the 16 wells.

20. A method according to any of claims 14-18, wherein the recognition sequence of the first and second set of oligonucleotide sequences consists of two bases and, prior to

performing the first step, the nucleic acid population is sub-divided into 256 wells, each well containing oligonucleotides from the first set of sequences having one of the 16 possible recognition sequences, and wherein in the second reaction oligonucleotides from the second set of sequences are added to each well, such that all possible combinations of the identity of the first and second set of oligonucleotide sequences and their order of addition to the well are represented in the 256 wells.

21. A method according to claim 19, wherein the contents of each pair of wells to which the same pair of oligonucleotide sequences were added but in a different order, are combined to give 10 different wells.

22. A method according to claim 20, wherein the contents of each pair of wells to which the same pair of oligonucleotide sequences were added but in a different order, are combined to give 136 different wells.

23. A method according to any preceding claim, wherein the oligonucleotide sequences have equalised melting temperatures.

24. A method according to claim 23, wherein the melting temperatures are equalised by incorporating one or more analogues of natural nucleotides into the oligonucleotide sequences, the analogues comprising base modifications, sugar modifications and/or backbone modifications.

25. A method according to any preceding claim, wherein the endonuclease is selected such that it cuts the nucleic acid at a site within the recognition site of the endonuclease.

26. A kit for categorising a nucleic acid, comprising one or more adaptors and one or more sets of oligonucleotide sequences, wherein the adaptors comprise nucleic acid having a double-stranded primer portion of a known sequence and a single-stranded portion of a pre-

determined length, either each single-stranded portion of each nucleic acid in the adaptors having the same pre-determined sequence or all possible sequences of the single-stranded portion being represented in the adaptors, and wherein each oligonucleotide sequence comprises a first sequence, a second sequence attached to the first sequence and a third sequence attached to the second sequence, in which the first sequence is complementary to the sequence of the primer portion of the adaptor, the second sequence is the same sequence as the single-stranded portion of the adaptors or all possible second sequences of the same length as the single-stranded portion of the adaptors are represented within the set of oligonucleotides, and the third sequence comprises a pre-determined recognition sequence.

27. A kit according to claim 26, wherein the recognition sequence consists of one base.

28. A kit according to claim 26, wherein in the recognition sequence consists of two or more bases.

29. A kit according to any of claims 26-28, wherein in any one group of oligonucleotides having the same recognition sequence, the third sequence consists of the recognition sequence and a pre-determined number of bases situated between the second sequence and the recognition sequence, all of the possible sequences of the pre-determined number of bases in the third sequence being represented in that group of oligonucleotides.

30. A kit according to any of claims 26-29, comprising two sets of oligonucleotide sequences, each of the oligonucleotides in one set being biotinylated.

31. A kit according to any of claims 26-29, comprising two sets of oligonucleotide sequences, each of the oligonucleotides in one set being covalently attached to a solid support.

32. A kit according to any of claims 26-31, additionally comprising an endonuclease.

33. A kit according to claim 32, wherein the endonuclease is selected such that when it is reacted with double-stranded nucleic acid, nucleic acids are produced each of which comprises a double-stranded portion.

34. A kit according to claim 33, wherein the endonuclease is selected such that the nucleic acids produced have a sticky end of a known common length extending from a terminal of the double-stranded portion, and wherein each sticky end of each nucleic acid in the nucleic acid population has the same known base sequence.

35. A kit according to claim 33, wherein the endonuclease is selected such that the nucleic acids produced have a sticky end of a known common length extending from a terminal of the double-stranded portion, and wherein the sticky ends of the nucleic acids in the nucleic acid population exhibit a plurality of different base sequences.

36. A kit according to any of claims 26-35, wherein the endonuclease is selected such that it cuts the nucleic acid at a site within the recognition site of the endonuclease.

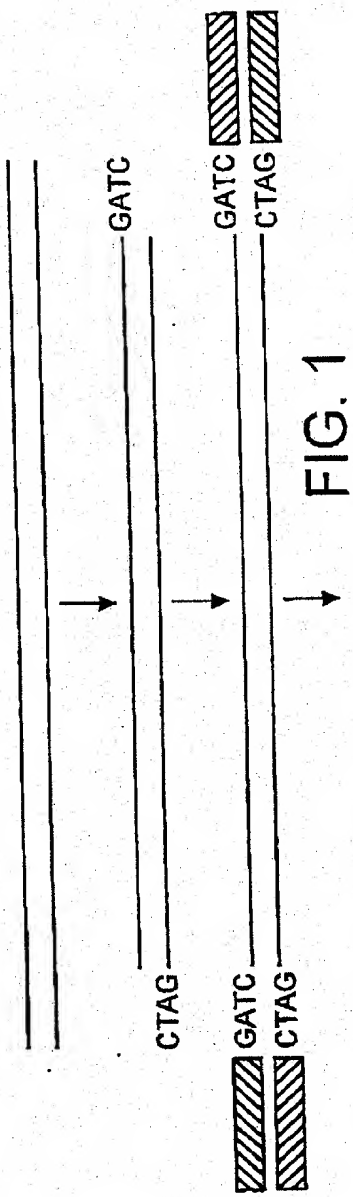

GATC
CTAG
GATC
CTAG
GATC
CTAG

FIG. 1

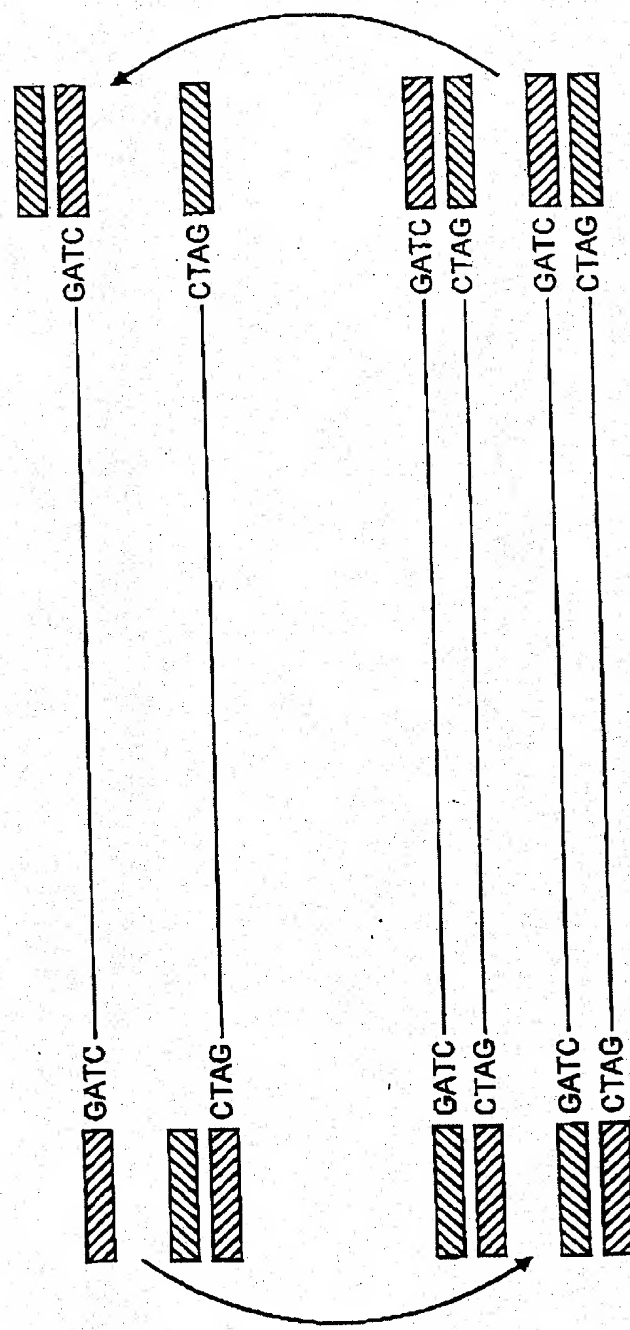

GATC
CTAG
GATC
CTAG
GATC
CTAG
GATC
CTAG
GATC
CTAG
GATC
CTAG

FIG. 2

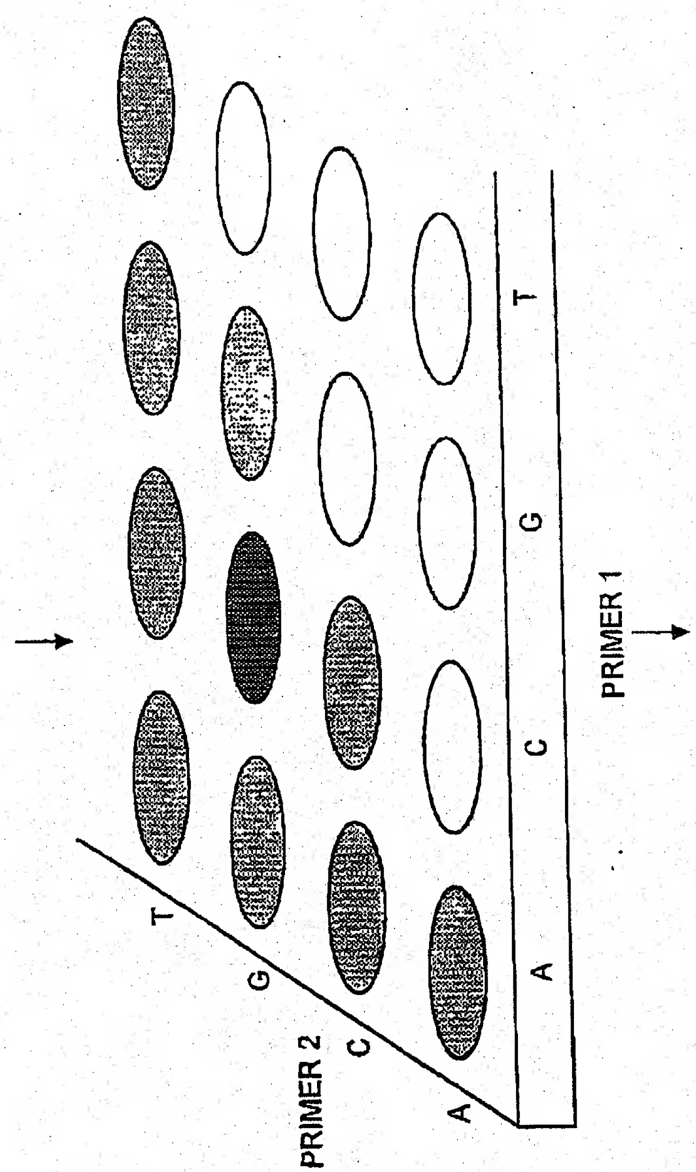


FIG. 3

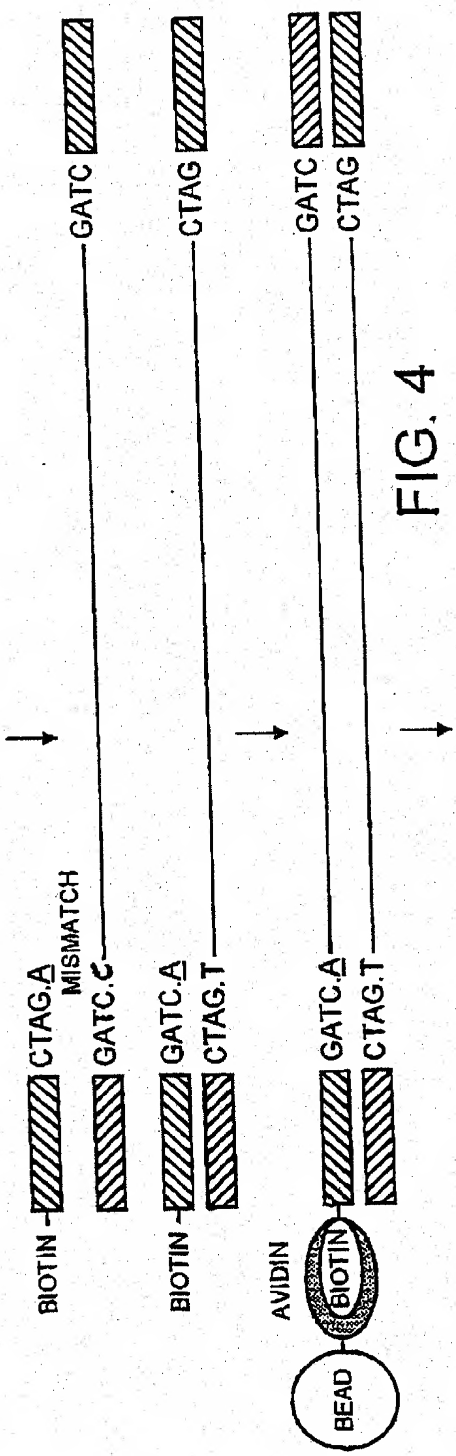

BIOTIN
CTAG.A
MISMATCH
GATC.C
GATC
GATC.A
CTAG.T
CTAG
AVIDIN
BIOTIN
BEAD

FIG. 4

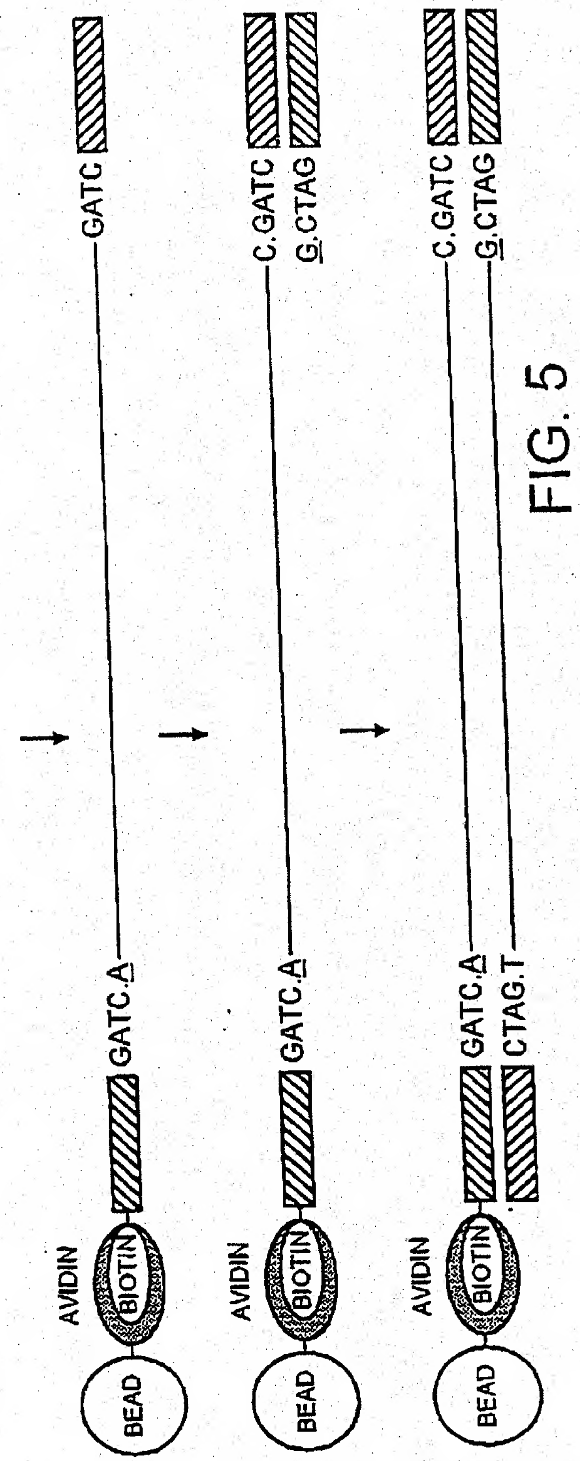

AVIDIN
BIOTIN
BEAD
GATC.A
GATC
C.GATC
G.CTAG
CTAG.T

FIG. 5

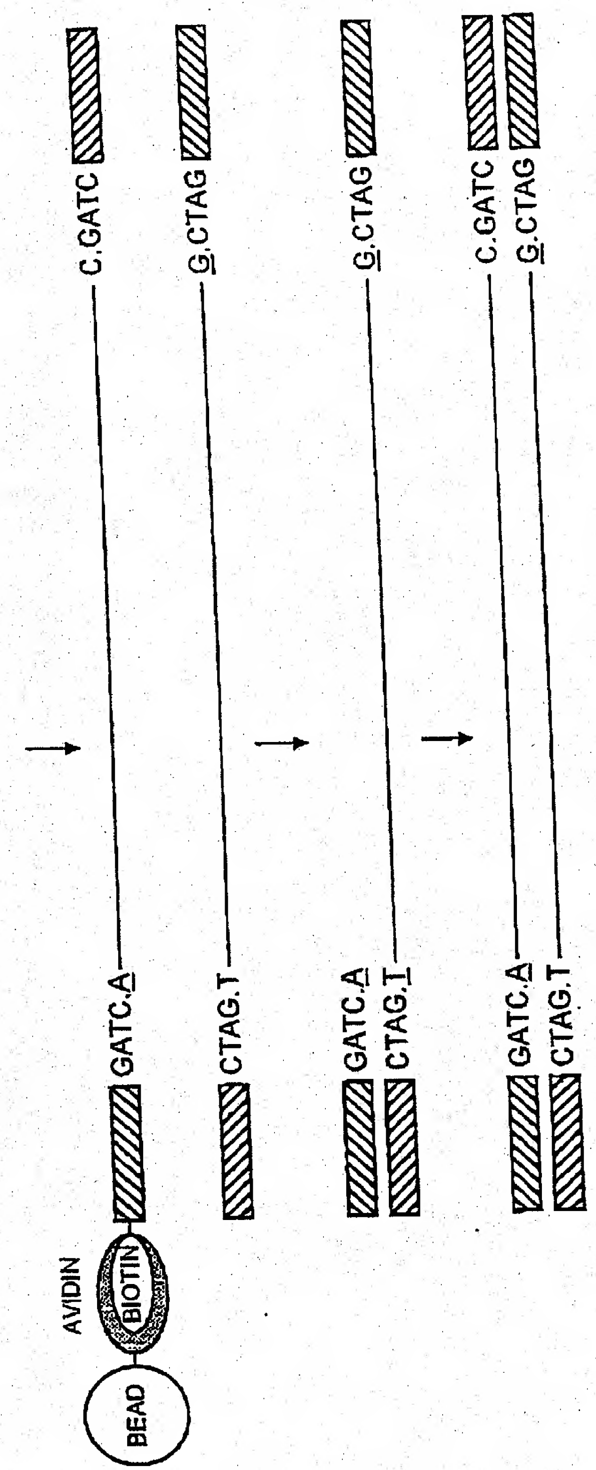

C.GATC
G.CTAG
G.CTAG
C.GATC
G.CTAG
GATC.A
CTAG.T
GATC.A
CTAG.T
GATC.A
CTAG.T
AVIDIN
BIOTIN
BEAD

FIG. 6A

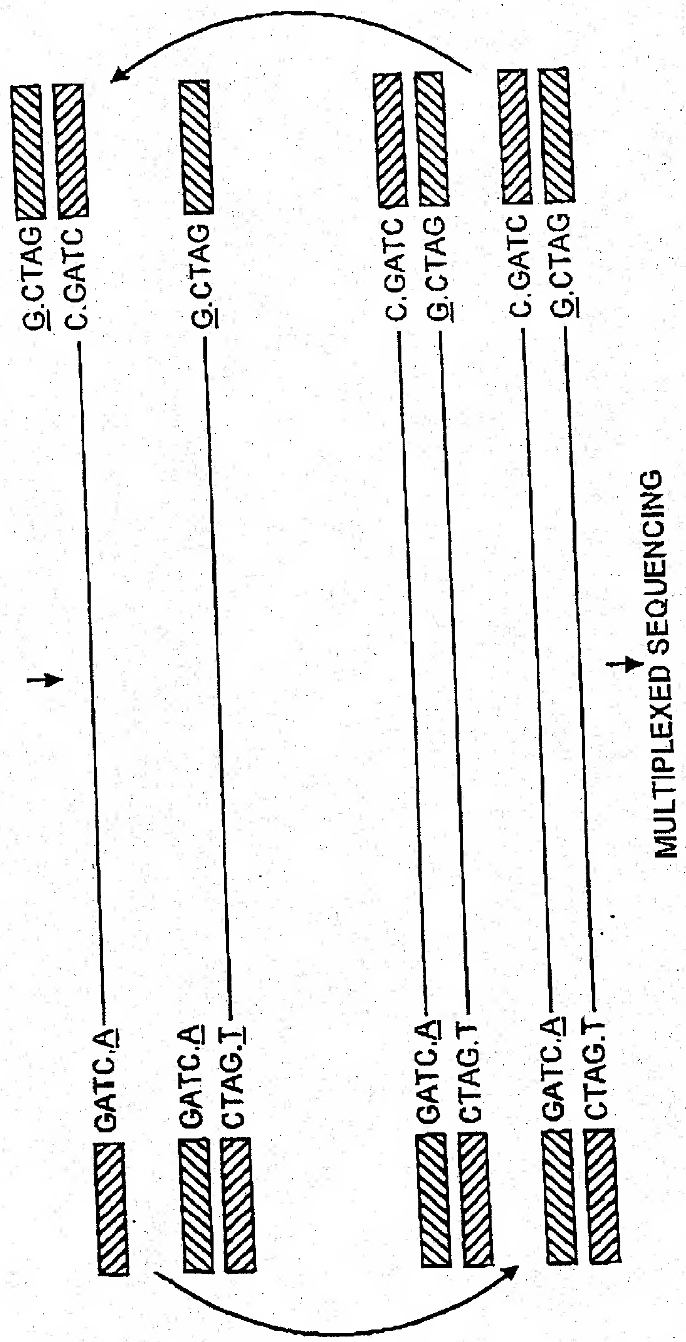

G.CTAG
C.GATC
G.CTAG
C.GATC
G.CTAG
C.GATC
G.CTAG
GATC.A
GATC.A
CTAG.T
GATC.A
CTAG.T
GATC.A
CTAG.T
MULTIPLEXED SEQUENCING

FIG. 6B

18

PATENT COOPERATION TREATY

REC'D 12 OCT 1999
WIPO PCT

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 86911/JND/CH	FOR FURTHER ACTION	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/GB98/02043	International filing date (day/month/year) 13/07/1998	Priority date (day/month/year) 11/07/1997
International Patent Classification (IPC) or national classification and IPC C12Q1/68		
Applicant BRAX GROUP LIMITED et al		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

These annexes consist of a total of 6 sheets.

3. This report contains indications relating to the following items:

I ☒ Basis of the report
II ☐ Priority
III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
IV ☐ Lack of unity of invention
V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
VI ☐ Certain documents cited
VII ☒ Certain defects in the international application
VIII ☐ Certain observations on the international application

Date of submission of the demand	Date of completion of this report
05/02/1999	07.10.99
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Tilkorn, A-C Telephone No. +49 89 2399 8688

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/GB98/02043

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-14	as originally filed			

Claims, No.:

1-36	as received on	06/09/1999	with letter of	03/09/1999

Drawings, sheets:

1/5-5/5	as originally filed			

2. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; *citations and explanations supporting such statement*

1. Statement

Novelty (N)	Yes:	Claims	1-25,29-31,35,36
	No:	Claims	26-28,32-34
Inventive step (IS)	Yes:	Claims	none
	No:	Claims	1-36
Industrial applicability (IA)	Yes:	Claims	1-36
	No:	Claims	none

2. Citations and explanations

see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

Section V:

The following documents are referred to in this communication:

D1: NUCLEIC ACIDS RESEARCH (1. May 1997) **25 (9)** pages 1854-1858
D2: WO 94 01582 A (cited in the application)

D1 describes a method for the amplification of specific DNA fragments including the ligation of adaptor molecules. The double stranded adaptors contain a universal primer sequence (M13 sequence), 5 nucleotides complementary to the protruding sequence produced by the restriction endonuclease (BclI) and 4 nucleotides which reside immediately adjacent to the restriction recognition sequence in the target molecule (p 1854 col 2 para 2; Fig 1). End-specific adaptors with pre-determined indexing sequences are disclosed (Fig 2A) and also the use of combinatorial adaptors (p 1857 col 1 para 1; Fig 2B) that contain a degenerate mixture of indexing sequences made in one oligodeoxynucleotide synthesis. Applications of this method are proposed, namely indexing of DNA populations and physical mapping and sequencing of whole genomes or sections of complex genomes (p 1857 col 2 para 3- p 1858 col 1 para 1).

D2 deals with a process for categorizing of nucleotide sequence populations (Title). D2 is cited and discussed in the description of the present application (appl.: p 1 para 4-p 2 para 2). The method of D2 includes the digestion of nucleic acid with an endonuclease, ligation of adaptors containing sequences that are complementary to the extending cleavage derived sequences (p 18 l 11-15) and separating adaptored products (p 8 para 4 - p 9 para 2). An embodiment of the method includes biontinylated adaptors which allow the capturing of the oligonucleotide on a solid phase (p 18 l 24- p 19 l 3; Example 1: p 37 l 5-p 38 l 16). Moreover the adaptor can comprise a known and selected sequence such that specifically adaptored molecules can be amplified by PCR using a primer complementary to the core sequence (p 19 l 3-7; Fig.1).

1. <u>Novelty</u>

1.1 **Claim 1** and the dependent **claims 2-25** are novel (Art 33(2)PCT), because in

none of the available documents all its technical features are disclosed:
- a method for categorizing nucleic acid comprising
- producing a nucleic acid population by action of an endonuclease on nucleic acid
- contacting the nucleic acid population with an oligonucleotide that contains a recognition sequence which enables a sequence specific hybridization with a double stranded part of some nucleic acids comprised in the population
- isolating nucleic acid by capturing the oligonucleotide on a solid phase

1.2 **Claim 26** relates to a kit containing the components to carry out the method disclosed in the application. It is not novel (Art 33(2)PCT) because D2 discloses a kit comprising adaptors, endonuclease e.g. *FokI* and oligonucleotide sequences used as PCR primers (D2: p 23 para 2; claims 29-34). The PCR primers comprise in some embodiments a sequence complementary to the core sequence of the adaptors (" first sequence") and may preferably extend by one or more specific extra bases into the adaptored fragment ("third sequence") (p 28 l 25- p 29 l 3). This implies, that the oligonucleotide sequences also contain the sequence of the single stranded portion of the adaptor ("second sequence"). Thus, these primers contain all the technical features of the oligonucleotide sequences claimed in claim 26. The same applies to claim **27-28, 32-34**.

1.3 **Claims 29-31, 35** and **36** which are dependent on claim 26 are novel, because they contain features that are not disclosed in D2 (cf. point 1.2 above).

2 Inventive Step:

2.1 **Claim 1** does not meet the requirements of Art 33(3)PCT for the following reasons:
D2, which is considered to represent the closest prior art, is distinguished from the subject-matter of claim 1 in that the oligonucleotide adaptors that correctly hybridize to the protruding single stranded portion of the nucleic acid are used to categorize the nucleic acid whereas according to claim 1 the oligonucleotide recognizes a sequence in the double stranded portion of the nucleic acid.

The problem to be solved over D2 can thus be regarded as how to provide a method for categorizing nucleic acid on a wider basis than the sequence of the

sticky ends (appl. p 2 para 2).

In order to solve the problem a skilled person would turn to D1, because D1 deals with indexing of DNA applicable for the accession and physical mapping of genomic DNA (abstract; p 1856 col 1 para 2; p 1857 col 2 para 3).
"Indexing" is understood to be a specific embodiment of "categorizing" as explained in D2 (p1 para 1; p 14 l 3-8; p 53 l 1-13). Indexing embraces the categorizing and the positioning of a marker (adaptor) at a predetermined site in a sequence.
The indexing sequence (D1: Fig 1) contained in the oligonucleotide adaptor hybridizes with a double stranded portion of the target nucleic acid. Thus, the adaptor of D1 has a predetermined recognition sequence that recognises a sequence in the double-stranded portion of the nucleic acid. By reacting the adaptor to a population of double stranded nucleic acids, that has been digested by an endonuclease, the adaptor hybridizes to nucleic acid molecules containing the recognition sequence in their double stranded portion. Hence, an adaptor according to D1 (e.g. Fig. 1) shows all technical features of the oligonucleotide sequence according to claim 1.

Thus, combining the general method of D2 with the adaptor oligonucleotide of D1 renders claim 1 obvious for a skilled person (Art 33(3)PCT).
The same applies to the dependent **claims 2,3, 6-8, 10, 11-13** as they do not contain an inventive concept per se.

2.2 **Claim 4** relates to a method in which an oligonucleotide adaptor is ligated first to the nucleic acids and then another oligonucleotide is used to hybridize with its recognition sequence on the double stranded portion of the nucleic acid before isolating the nucleic acid which hybridizes correctly with the oligonucleotide sequence.
In D2, which is considered to represent the closest prior art, adaptor oligonucleotides are described that contain a known and selected sequence such that nucleic acids linked to an adaptor can be amplified by PCR using a primer complementary to the core sequence (p 19 l 3-7). These adaptors appear to contain the same technical features as the adaptors of the present claim 4.
D2 is distinguished from the subject-matter of claim 4 in that in D2 the nucleic

acids that correctly hybridize with the adaptor oligonucleotide are isolated whereas in the method according to claim 4 the nucleic acids are isolated that correctly hybridize to another oligonucleotide. Moreover, the adaptor disclosed in D2 hybridizes to a single stranded portion of the nucleic acid (see point 2 above).
The problem to be solved over D2 can also be regarded as how to provide method for categorizing nucleic acid on a wider basis than the sequence of the sticky ends (appl. p 2 para 2).
The method of D2 including the amplification of the adaptored nucleic acids by primers that hybridize with a complementary portion of the adaptor (D2: p 19 l 3-7) is suitable for categorizing nucleic acids on a wider basis than the sequence of the sticky ends. In D2 the general technique is disclosed for the isolation of adaptored nucleic acids (p 18 l 24- p 19 l 3 ; p 28 l 12-p 29 l 3). In addition the physical separation of the initial adaptored nucleic acids is discussed and it is stated, that it is not strictly necessary, if a PCR-type process using selected primers is employed (p 28 l 17-19). Moreover, subsets of sequences obtained by the method of D2 can themselves be immobilized by standard techniques for further analysis (p 32 l 26-p 33 l 2).
Thus, a skilled person would be able to isolate the amplification products by common technology without being inventive (Art 33(3)PCT). The same applies to the dependent **claims 5,9-25** as they do not contain an inventive concept per se.

2.3 Although **claims 29-31, 35** and **36** which are dependent on claim 26 contain new technical features they do not appear to be inventive as they appear to be simple modifications that can be achieved using common technology by a skilled person (Art 33(3)PCT).

Section VII:

- Contrary to the requirements of Rule 5.1(a)(ii) PCT, the relevant background art disclosed in the document D1 is not mentioned in the description, nor is this document identified therein.

- In order for the application to be self-contained, patent application numbers should have been replaced by the corresponding publication numbers (e.g. p 1 para 4; p 6 l 1).

PATENT COOPERATION TREATY

From the INTERNATIONAL BUREAU

To:

DANIELS, Jeffrey, Nicholas
Page White & Farrer
54 Doughty Street
London WC1N 2LS
ROYAUME-UNI

PCT

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

Date of mailing (day/month/year)
22 April 1999 (22.04.99)

Applicant's or agent's file reference
86911/JND/CH

IMPORTANT NOTIFICATION

International application No.
PCT/GB98/02043

International filing date (day/month/year)
13 July 1998 (13.07.98)

1. The following indications appeared on record concerning:

[X] the applicant [ ] the inventor [ ] the agent [ ] the common representative

Name and Address	State of Nationality	State of Residence
BRAX GENOMICS LIMITED 13 Station Road Cambridge CB1 2JB United Kingdom	GB	GB
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

[ ] the person [X] the name [ ] the address [ ] the nationality [ ] the residence

Name and Address	State of Nationality	State of Residence
BRAX GROUP LIMITED 13 Station Road Cambridge CB1 2JB United Kingdom	GB	GB
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

[X] the receiving Office
[ ] the International Searching Authority
[X] the International Preliminary Examining Authority
[ ] the designated Offices concerned
[X] the elected Offices concerned
[ ] other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

Nicola Wolff

Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATTY

From the INTERNATIONAL BUREAU

PCT

To:

DANIELS, Jeffrey, Nicholas
Page White & Farrer
54 Doughty Street
London WC1N 2LS
ROYAUME-UNI

28 APR 1999

NOTIFICATION OF THE RECORDING OF A CHANGE

(PCT Rule 92bis.1 and Administrative Instructions, Section 422)

Date of mailing (day/month/year)
22 April 1999 (22.04.99)

Applicant's or agent's file reference
86911/JND/CH

IMPORTANT NOTIFICATION

International application No.
PCT/GB98/02043

International filing date (day/month/year)
13 July 1998 (13.07.98)

1. The following indications appeared on record concerning:

[X] the applicant [ ] the inventor [ ] the agent [ ] the common representative

Name and Address	State of Nationality	State of Residence
BRAX GENOMICS LIMITED 13 Station Road Cambridge CB1 2JB United Kingdom	GB	GB
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

2. The International Bureau hereby notifies the applicant that the following change has been recorded concerning:

[ ] the person [X] the name [ ] the address [ ] the nationality [ ] the residence

Name and Address	State of Nationality	State of Residence
BRAX GROUP LIMITED 13 Station Road Cambridge CB1 2JB United Kingdom	GB	GB
	Telephone No.	
	Facsimile No.	
	Teleprinter No.	

3. Further observations, if necessary:

4. A copy of this notification has been sent to:

[X] the receiving Office
[ ] the International Searching Authority
[X] the International Preliminary Examining Authority
[ ] the designated Offices concerned
[X] the elected Offices concerned
[ ] other:

The International Bureau of WIPO
34, chemin des Colombettes
1211 Geneva 20, Switzerland

Facsimile No.: (41-22) 740.14.35

Authorized officer

NWolff

Nicola Wolff

Telephone No.: (41-22) 338.83.38

PATENT COOPERATION TREATY

RECEIVED
11 OCT 1999

From the
INTERNATIONAL PRELIMINARY EXAMINING AUTHORITY

To:

DANIELS, Jeffrey N.
PAGE WHITE & FARRER
54 Doughty Street
LONDON WC1N 2LS
GRANDE BRETAGNE

PCT

NOTIFICATION OF TRANSMITTAL OF THE INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Rule 71.1)

Date of mailing (*day/month/year*)	07.10.99
Applicant's or agent's file reference 86911/JND/CH	**IMPORTANT NOTIFICATION**

International application No.	International filing date (*day/month/year*)	Priority date (*day/month/year*)
PCT/GB98/02043	13/07/1998	11/07/1997

Applicant
BRAX GROUP LIMITED et al

1. The applicant is hereby notified that this International Preliminary Examining Authority transmits herewith the international preliminary examination report and its annexes, if any, established on the international application.

2. A copy of the report and its annexes, if any, is being transmitted to the International Bureau for communication to all the elected Offices.

3. Where required by any of the elected Offices, the International Bureau will prepare an English translation of the report (but not of any annexes) and will transmit such translation to those Offices.

4. **REMINDER**

The applicant must enter the national phase before each elected Office by performing certain acts (filing translations and paying national fees) within 30 months from the priority date (or later in some Offices) (Article 39(1)) (see also the reminder sent by the International Bureau with Form PCT/IB/301).

Where a translation of the international application must be furnished to an elected Office, that translation must contain a translation of any annexes to the international preliminary examination report. It is the applicant's responsibility to prepare and furnish such translation directly to each elected Office concerned.

For further details on the applicable time limits and requirements of the elected Offices, see Volume II of the PCT Applicant's Guide.

Name and mailing address of the IPEA/	Authorized officer
European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Digiusto, M Tel.+49 89 2399-8162

PATENT COOPERATION TREATY

PCT

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

(PCT Article 36 and Rule 70)

Applicant's or agent's file reference 86911/JND/CH	**FOR FURTHER ACTION**	See Notification of Transmittal of International Preliminary Examination Report (Form PCT/IPEA/416)
International application No. PCT/GB98/02043	International filing date *(day/month/year)* 13/07/1998	Priority date *(day/month/year)* 11/07/1997
International Patent Classification (IPC) or national classification and IPC C12Q1/68		
Applicant BRAX GROUP LIMITED et al		

1. This international preliminary examination report has been prepared by this International Preliminary Examining Authority and is transmitted to the applicant according to Article 36.

2. This REPORT consists of a total of 7 sheets, including this cover sheet.

   ☒ This report is also accompanied by ANNEXES, i.e. sheets of the description, claims and/or drawings which have been amended and are the basis for this report and/or sheets containing rectifications made before this Authority (see Rule 70.16 and Section 607 of the Administrative Instructions under the PCT).

   These annexes consist of a total of 6 sheets.

3. This report contains indications relating to the following items:

   I ☒ Basis of the report
   II ☐ Priority
   III ☐ Non-establishment of opinion with regard to novelty, inventive step and industrial applicability
   IV ☐ Lack of unity of invention
   V ☒ Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations suporting such statement
   VI ☐ Certain documents cited
   VII ☒ Certain defects in the international application
   VIII ☐ Certain observations on the international application

Date of submission of the demand 05/02/1999	Date of completion of this report 07.10.99
Name and mailing address of the international preliminary examining authority: European Patent Office D-80298 Munich Tel. +49 89 2399 - 0 Tx: 523656 epmu d Fax: +49 89 2399 - 4465	Authorized officer Tilkorn, A-C Telephone No. +49 89 2399 8688

INTERNATIONAL PRELIMINARY EXAMINATION REPORT

International application No. PCT/GB98/02043

I. Basis of the report

1. This report has been drawn on the basis of (*substitute sheets which have been furnished to the receiving Office in response to an invitation under Article 14 are referred to in this report as "originally filed" and are not annexed to the report since they do not contain amendments.*):

Description, pages:

1-14	as originally filed			

Claims, No.:

1-36	as received on	06/09/1999	with letter of	03/09/1999

Drawings, sheets:

1/5-5/5	as originally filed			

2. The amendments have resulted in the cancellation of:

- ☐ the description, pages:
- ☐ the claims, Nos.:
- ☐ the drawings, sheets:

3. ☐ This report has been established as if (some of) the amendments had not been made, since they have been considered to go beyond the disclosure as filed (Rule 70.2(c)):

4. Additional observations, if necessary:

V. Reasoned statement under Article 35(2) with regard to novelty, inventive step or industrial applicability; citations and explanations supporting such statement

1. Statement

Novelty (N)	Yes:	Claims	1-25,29-31,35,36
	No:	Claims	26-28,32-34
Inventive step (IS)	Yes:	Claims	none
	No:	Claims	1-36
Industrial applicability (IA)	Yes:	Claims	1-36
	No:	Claims	none

2. Citations and explanations

see separate sheet

VII. Certain defects in the international application

The following defects in the form or contents of the international application have been noted:

see separate sheet

Section V:

The following documents are referred to in this communication:

D1: NUCLEIC ACIDS RESEARCH (1. May 1997) **25 (9)** pages 1854-1858
D2: WO 94 01582 A (cited in the application)

D1 describes a method for the amplification of specific DNA fragments including the ligation of adaptor molecules. The double stranded adaptors contain a universal primer sequence (M13 sequence), 5 nucleotides complementary to the protruding sequence produced by the restriction endonuclease (BclI) and 4 nucleotides which reside immediately adjacent to the restriction recognition sequence in the target molecule (p 1854 col 2 para 2; Fig 1). End-specific adaptors with pre-determined indexing sequences are disclosed (Fig 2A) and also the use of combinatorial adaptors (p 1857 col 1 para 1; Fig 2B) that contain a degenerate mixture of indexing sequences made in one oligodeoxynucleotide synthesis. Applications of this method are proposed, namely indexing of DNA populations and physical mapping and sequencing of whole genomes or sections of complex genomes (p 1857 col 2 para 3- p 1858 col 1 para 1).

D2 deals with a process for categorizing of nucleotide sequence populations (Title). D2 is cited and discussed in the description of the present application (appl.: p 1 para 4-p 2 para 2). The method of D2 includes the digestion of nucleic acid with an endonuclease, ligation of adaptors containing sequences that are complementary to the extending cleavage derived sequences (p 18 l 11-15) and separating adaptored products (p 8 para 4 - p 9 para 2). An embodiment of the method includes biontinylated adaptors which allow the capturing of the oligonucleotide on a solid phase (p 18 l 24- p 19 l 3; Example 1; p 37 l 5-p 38 l 16). Moreover the adaptor can comprise a known and selected sequence such that specifically adaptored molecules can be amplified by PCR using a primer complementary to the core sequence (p 19 l 3-7; Fig.1).

1. <u>Novelty</u>

1.1 **Claim 1** and the dependent **claims 2-25** are novel (Art 33(2)PCT), because in

none of the available documents all its technical features are disclosed:
- a method for categorizing nucleic acid comprising
- producing a nucleic acid population by action of an endonuclease on nucleic acid
- contacting the nucleic acid population with an oligonucleotide that contains a recognition sequence which enables a sequence specific hybridization with a double stranded part of some nucleic acids comprised in the population
- isolating nucleic acid by capturing the oligonucleotide on a solid phase

1.2 **Claim 26** relates to a kit containing the components to carry out the method disclosed in the application. It is not novel (Art 33(2)PCT) because D2 discloses a kit comprising adaptors, endonuclease e.g. *FokI* and oligonucleotide sequences used as PCR primers (D2: p 23 para 2; claims 29-34). The PCR primers comprise in some embodiments a sequence complementary to the core sequence of the adaptors (" first sequence") and may preferably extend by one or more specific extra bases into the adaptored fragment ("third sequence") (p 28 l 25- p 29 l 3). This implies, that the oligonucleotide sequences also contain the sequence of the single stranded portion of the adaptor ("second sequence"). Thus, these primers contain all the technical features of the oligonucleotide sequences claimed in claim 26. The same applies to claim **27-28, 32-34**.

1.3 **Claims 29-31, 35** and **36** which are dependent on claim 26 are novel, because they contain features that are not disclosed in D2 (cf. point 1.2 above).

2 <u>Inventive Step:</u>

2.1 **Claim 1** does not meet the requirements of Art 33(3)PCT for the following reasons:
D2, which is considered to represent the closest prior art, is distinguished from the subject-matter of claim 1 in that the oligonucleotide adaptors that correctly hybridize to the protruding <u>single stranded</u> portion of the nucleic acid are used to categorize the nucleic acid whereas according to claim 1 the oligonucleotide recognizes a sequence in the double stranded portion of the nucleic acid.

The problem to be solved over D2 can thus be regarded as how to provide a method for categorizing nucleic acid on a wider basis than the sequence of the

sticky ends (appl. p 2 para 2).

In order to solve the problem a skilled person would turn to D1, because D1 deals with indexing of DNA applicable for the accession and physical mapping of genomic DNA (abstract; p 1856 col 1 para 2; p 1857 col 2 para 3).
"Indexing" is understood to be a specific embodiment of "categorizing" as explained in D2 (p1 para 1; p 14 l 3-8; p 53 l 1-13). Indexing embraces the categorizing and the positioning of a marker (adaptor) at a predetermined site in a sequence.
The indexing sequence (D1: Fig 1) contained in the oligonucleotide adaptor hybridizes with a double stranded portion of the target nucleic acid. Thus, the adaptor of D1 has a predetermined recognition sequence that recognises a sequence in the double-stranded portion of the nucleic acid. By reacting the adaptor to a population of double stranded nucleic acids, that has been digested by an endonuclease, the adaptor hybridizes to nucleic acid molecules containing the recognition sequence in their double stranded portion. Hence, an adaptor according to D1 (e.g. Fig. 1) shows all technical features of the oligonucleotide sequence according to claim 1.

Thus, combining the general method of D2 with the adaptor oligonucleotide of D1 renders claim 1 obvious for a skilled person (Art 33(3)PCT).
The same applies to the dependent **claims 2,3, 6-8, 10, 11-13** as they do not contain an inventive concept per se.

2.2 **Claim 4** relates to a method in which an oligonucleotide adaptor is ligated first to the nucleic acids and then another oligonucleotide is used to hybridize with its recognition sequence on the double stranded portion of the nucleic acid before isolating the nucleic acid which hybridizes correctly with the oligonucleotide sequence.
In D2, which is considered to represent the closest prior art, adaptor oligonucleotides are described that contain a known and selected sequence such that nucleic acids linked to an adaptor can be amplified by PCR using a primer complementary to the core sequence (p 19 l 3-7). These adaptors appear to contain the same technical features as the adaptors of the present claim 4.
D2 is distinguished from the subject-matter of claim 4 in that in D2 the nucleic

acids that correctly hybridize with the adaptor oligonucleotide are isolated whereas in the method according to claim 4 the nucleic acids are isolated that correctly hybridize to another oligonucleotide. Moreover, the adaptor disclosed in D2 hybridizes to a single stranded portion of the nucleic acid (see point 2 above). The problem to be solved over D2 can also be regarded as how to provide method for categorizing nucleic acid on a wider basis than the sequence of the sticky ends (appl. p 2 para 2).
The method of D2 including the amplification of the adaptored nucleic acids by primers that hybridize with a complementary portion of the adaptor (D2: p 19 l 3-7) is suitable for categorizing nucleic acids on a wider basis than the sequence of the sticky ends. In D2 the general technique is disclosed for the isolation of adaptored nucleic acids (p 18 l 24- p 19 l 3 ; p 28 l 12-p 29 l 3). In addition the physical separation of the initial adaptored nucleic acids is discussed and it is stated, that it is not strictly necessary, if a PCR-type process using selected primers is employed (p 28 l 17-19). Moreover, subsets of sequences obtained by the method of D2 can themselves be immobilized by standard techniques for further analysis (p 32 l 26-p 33 l 2).
Thus, a skilled person would be able to isolate the amplification products by common technology without being inventive (Art 33(3)PCT). The same applies to the dependent **claims 5,9-25** as they do not contain an inventive concept per se.

2.3 Although **claims 29-31, 35** and **36** which are dependent on claim 26 contain new technical features they do not appear to be inventive as they appear to be simple modifications that can be achieved using common technology by a skilled person (Art 33(3)PCT).

Section VII:

- Contrary to the requirements of Rule 5.1(a)(ii) PCT, the relevant background art disclosed in the document D1 is not mentioned in the description, nor is this document identified therein.

- In order for the application to be self-contained, patent application numbers should have been replaced by the corresponding publication numbers (e.g. p 1 para 4; p 6 l 1).